HARRY G. FRANKFURT

La importancia de lo que nos preocupa

ENSAYOS FILOSÓFICOS

conocimiento

katz

Harry G. Frankfurt
(1929, Pennsylvania, Estados Unidos)
Obtuvo el doctorado en filosofía
en la Universidad John Hopkins en 1949.
Fue profesor en la Universidad de Ohio,
en la Universidad Estatal de Nueva York
(SUNY) y en la Universidad Rockefeller.
Desde 1976 enseñó en Yale, cuyo Departamento
de Filosofía dirigió entre 1978 y 1987.
Desde 1990 fue profesor en la Universidad
de Princeton, de la que es actualmente profesor
emérito. Frankfurt mismo ha declarado que
su interés no es la moralidad en el sentido
estricto del término, sino más bien
las consecuencias éticas del deseo, el amor,
el cuidado... Sus análisis abren nuevas maneras
de pensar la libertad, la responsabilidad,
el altruismo y el amor.

La importancia de lo que nos preocupa

Del mismo autor

Las razones del amor, Barcelona, 2004
Necessity, volition, and love, Cambridge, 1998 (edición española en preparación por Katz Editores)
On Bullshit, Princeton, New Jersey, 2005 (la versión original de este breve ensayo fue publicada originalmente en *La importancia de lo que nos preocupa)*
Demons, dreamers and madmen: the defense of reason in Descartes' meditations, Nueva York y Londres, 1970
Freiheit und Selbstbestimmung, Berlín, 2001

Harry G. Frankfurt

La importancia de lo que nos preocupa

Ensayos filosóficos

Traducido por Verónica Inés Weinstabl
y Servanda María de Hagen

katz conocimiento

Frankfurt, Harry G.
La importancia de lo que nos preocupa : ensayos filosóficos
- la ed. - Buenos Aires : Katz, 2006.
274 p. ; 23x15 cm.

Traducido por: Verónica Inés Weinstabl y Servanda María de Hagen

ISBN 987-1283-04-0

1. Filosofía Occidental. I. Weinstabl, Verónica Inés y De Hagen, Servanda María, trad. II. Título
CDD 190

Primera edición, 2006

© Katz Editores
Sinclair 2949, 5º B
1428, Buenos Aires
www.katzeditores.com

Título de la edición original: *The Importance of What We Care About. Philosophical Essays*, publicado por Press Syndicate of the University of Cambridge
© Cambridge University Press
Nueva York, 1988

ISBN Argentina: 987-1283-04-0
ISBN España: 84-609-8358-7

CREATIVE COMMONS

Diseño de colección: tholön kunst
Impreso en la Argentina por Latingráfica S. R. L.

Hecho el depósito que marca la ley 11.723.

Índice

7 Prefacio

11 Posibilidades alternativas y responsabilidad moral
25 La libertad de la voluntad y el concepto de persona
47 Coacción y responsabilidad moral
75 Tres conceptos de *acción libre*
89 Identificación y externalidad
105 El problema de la acción
119 La importancia de lo que nos preocupa
139 De qué somos moralmente responsables
153 Necesidad y deseo
171 Sobre el concepto de *bullshit*
195 La igualdad como ideal moral
229 Identificación e incondicionalidad
253 La racionalidad y lo impensable

272 Fuentes

Prefacio

Los ensayos aquí compilados tendrán que hablar por sí mismos, por supuesto, y no tiene mucho sentido que ahora intente resumir o parafrasear lo que creo que dice cada uno de ellos. Sin duda, sería más adecuado que me refiriera, en forma sucinta pero integral, a las ambiciones o a los temas filosóficos generales (suponiendo que los haya) en los que convergen los ensayos y mediante los cuales –a pesar de que los textos han sido concebidos y compuestos por separado a lo largo de veinte años– en cierto modo se unifican. Estoy seguro de que la convergencia existe, teniendo en cuenta lo que sé acerca de las preocupaciones personales a las que responde mi trabajo filosófico y su carácter ineludible en mi vida. Sin embargo, esto no quiere decir, a mi pesar, que comprenda lo suficientemente bien lo que he hecho para poder brindar una versión clara y directa de ello. Junto con las observaciones fragmentarias que puedo ofrecer aquí, debo admitir con cierta inquietud lo improbable de que mis pensamientos preliminares sean más perspicaces en su objetivo o más firmes en sus conocimientos que aquellos que aparecen en los ensayos mismos.

Varios de los ensayos tratan bastante detenidamente, de una manera u otra, temas relacionados con la naturaleza y las condiciones de la responsabilidad moral. Esta insistencia puede ser algo engañosa. La moralidad, como la entiendo, tiene que ver, en particular, con la forma en que debemos conducirnos en nuestra relación con los demás. Algunos de los ensayos se refieren a ello. Sin embargo, mi atención filosófica, en su mayor parte, ha sido guiada no tanto por un interés en las cuestiones sobre la moralidad como

por una preocupación con los temas que pertenecen, más apropiadamente, a la metafísica o a la filosofía de la mente: por ejemplo, cómo hemos de conceptualizarnos como personas, y qué define las identidades que logramos.

Me inclino a pensar que la definición de estos temas depende, en gran medida, del desarrollo de caracterizaciones adecuadas de la libertad y los ideales personales y no, particularmente, de la formulación de una teoría de obligaciones y derechos. Sin duda, la filosofía moral cumple una función en el intento de comprender esos asuntos, pero, en mi opinión, su función es más limitada de lo que a menudo se supone. En gran parte, los ideales a los cuales una persona dedica libremente su vida no son en forma exclusiva –ni siquiera prioritariamente– ideales morales. He hecho algunos intentos, en el ensayo cuyo título tomé para el libro y en algunos que escribí con posterioridad, en pos del desarrollo de una teoría de los ideales. Éste es, hasta un punto sorprendente y que debemos lamentar, un tema algo descuidado, acerca del cual quisiera poder decir más.

En lo que respecta a la libertad, es cierto, desde ya, que comúnmente se la entiende como una condición necesaria de la responsabilidad moral. Más aun, con frecuencia se investiga la naturaleza de la libertad –como yo mismo lo intenté algunas veces– a la luz de esa relación particular. Por otra parte, no creo que nuestra preocupación por ser libres se deba principalmente al respeto por la responsabilidad moral. Tampoco estoy convencido de que sea posible dilucidar lo que significa para nosotros ser libres sólo si comenzamos por interpretar que la libertad es una condición específica que requiere el ser moralmente responsable.

Las consideraciones que mis ensayos proponen en relación con estos asuntos son, en esencia, estructurales más que históricas. Comprender qué es una persona, ya sea como entidad de cierto tipo genérico, ya como individuo, es diferente de comprender cómo llegó a ser lo que es. Ambos tipos de preguntas son importantes e interesantes y, en cierta manera, están estrechamente relacionadas. Cada uno de nosotros es, en forma indiscutible, una criatura de la historia. Sin el pasado, no seríamos nada; y sólo cuando una persona reconoce que ella misma es, en su totalidad, el

producto de contingencias históricas –biológicas, sociales y personales–, puede identificar y comprender su propia naturaleza. Sin embargo, la empresa más auténticamente filosófica no es averiguar qué nos produjo, sino tratar de identificar y comprender en qué nos hemos convertido. Esta insistencia en tratar de ver las cosas claramente por lo que son y no en términos de otras cosas es, por cierto, lo que vincula con más facilidad los temas de mis ensayos sobre el *bullshit** y el igualitarismo con los temas del resto de los capítulos.

En general, el enfoque que adopto para intentar comprender lo que somos es considerar la estructura y la constitución del yo. El énfasis que pongo en esto se encuentra, principalmente, en la voluntad. Con frecuencia, la razón ha sido considerada la característica más distintiva de la naturaleza humana y la más claramente definitiva. Sin embargo, creo que la volición se relaciona en forma más estrecha que la razón con nuestra experiencia de nosotros mismos y con los problemas de nuestras vidas que nos resultan más urgentes. Así, es la facultad más personal y más íntima e, incluso, puede ser también la más integral. En cualquier caso, los temas principales que he intentado explorar son la organización interna de la voluntad y lo que ello implica para nosotros.

En el siglo XVII, se estableció el mecanismo como cosmovisión dominante de nuestra cultura. Desde esa época ha llegado a parecer obvio que, o bien las referencias a las causas finales son totalmente ilícitas, o bien que no son más que formas de hablar convenientes destinadas a evitar formulaciones más toscas (aunque estrictamente sean más precisas) en términos de causalidad eficiente. En el siglo XVIII, la noción de una causa eficiente fue, en sí misma, eviscerada por una crítica devastadora de la idea del poder inherente. Estos persuasivos acontecimientos filosóficos han tornado difícil la tarea de ofrecer una buena explicación de la diferencia entre ser activo y pasivo, ya que, si se considera que las cosas no tienen ni objetivos ni poderes, ¿de qué modo sería posible entenderlas como

* Tal como se advierte en el capítulo 10, se ha mantenido en inglés el término *bullshit*, que se puede traducir, según el contexto, como "tonterías", "mentiras", "sandeces", "joder", o con otros vocablos similares o emparentados. [N. de T.]

activas? No obstante, el papel de la distinción entre activo y pasivo en la vida humana es omnipresente y profundo. La diferencia entre pasividad y actividad es inescindible del hecho de que existimos como seres y agentes, y no sólo como escenarios en los que, por casualidad, ocurren ciertos acontecimientos.

Aunque no me he ocupado de analizar esta distinción como tal, figura de diversas maneras en los numerosos análisis de la necesidad y la voluntad que transmiten otra preocupación temática de mis ensayos. Me parece que descubrir qué somos es, fundamentalmente –aunque sólo quizá entre otras cuestiones–, descubrir qué debemos ser. Y, por ello, en la medida en que una persona se identifica por su voluntad, es cuestión de descubrir lo que no podemos evitar desear o lo que no podemos lograr desear. La noción de que la necesidad no debilita en forma inevitable la autonomía es conocida y ampliamente aceptada. Sin embargo, la necesidad no sólo es compatible con la autonomía; en ciertos aspectos, le es esencial. Debe haber límites a nuestra libertad si hemos de tener una realidad personal suficiente para ejercer la auténtica autonomía. Lo que no tiene límites, no tiene forma. Del mismo modo, una persona no puede tener una identidad o una naturaleza esencial como agente a menos que esté obligada con respecto a esa característica de sí misma –es decir, la voluntad–, cuya forma coincide, fielmente, con lo que es esa persona, y que también lo da a conocer. Varios de mis ensayos –entre ellos el último, que hasta ahora no había sido publicado– se refieren a los intentos de arrojar luz sobre la creación de estas necesidades y sobre la importancia que tienen para nosotros.

1
Posibilidades alternativas y responsabilidad moral

Un principio al que llamaré el *principio de las posibilidades alternativas* ha desempeñado un papel dominante en casi todas las investigaciones recientes acerca del problema del libre albedrío. Este principio afirma que una persona es moralmente responsable de lo que ha hecho sólo en el caso de que hubiera podido comportarse de otra manera. Su significado exacto es un tema de controversia, en particular con respecto a la cuestión de si alguien que lo acepta queda comprometido a creer que la responsabilidad moral y el determinismo son incompatibles. Sin embargo, prácticamente nadie parece inclinado a negar o siquiera a cuestionar que el principio de las posibilidades alternativas (interpretado de una u otra manera) sea verdad. En general, siempre pareció tan abrumadoramente razonable que algunos filósofos han llegado a caracterizarlo como una verdad *a priori*. Es evidente que las personas cuyas concepciones de libre albedrío o de responsabilidad moral están radicalmente enfrentadas encuentran en él un terreno común firme y conveniente sobre el que pueden adoptar con provecho sus posturas opuestas.

Pero el principio de las posibilidades alternativas es falso. Una persona bien puede ser moralmente responsable de lo que hizo aunque no haya podido comportarse de otra manera. La verosimilitud del principio es una ilusión, que puede desvanecerse si centramos la atención en los fenómenos morales pertinentes.

I

Cuando se busca ilustrar el principio de las posibilidades alternativas, es muy natural pensar en situaciones en las que las mismas circunstancias provocan que una persona haga algo y, a la vez, hacen imposible que evite hacerlo. Éstas incluyen, por ejemplo, situaciones en las que una persona es coaccionada a hacer algo o es impelida de actuar por una sugestión hipnótica, o en las que un impulso interno la lleva a actuar como lo hace. En situaciones de esta índole, hay circunstancias que impiden que la persona actúe de otra manera, y estas mismas circunstancias también sirven para provocar que haga lo que hace.

No obstante, puede haber circunstancias que constituyen condiciones suficientes para que alguien lleve a cabo una acción y que, por lo tanto, impiden que la persona actúe de otra forma, pero que, en realidad, no obligan a la persona a actuar ni a producir su acción de ninguna manera. Una persona puede hacer algo en circunstancias que no le dejan más alternativa que hacerlo, sin que estas circunstancias en realidad la muevan o la conduzcan a hacerlo, sin que desempeñen ningún papel, en efecto, en provocar que actúe como lo hace.

Un examen de situaciones caracterizadas por circunstancias de esta clase arroja dudas, creo yo, sobre la relevancia del hecho de que una persona que ha actuado de una manera no podría haber actuado de otra, respecto de cuestiones de responsabilidad moral. Propongo desarrollar algunos ejemplos de esta clase en el contexto de un debate sobre la coacción y sugerir que nuestras intuiciones morales con respecto a estos ejemplos tienden a desconfirmar el principio de las posibilidades alternativas. Luego discutiré el principio en términos más generales, explicaré lo que pienso que son sus errores y describiré brevemente y sin argumento cómo podría revisarse en forma apropiada.

II

En general, estamos de acuerdo en que una persona que ha sido coaccionada a hacer algo no lo hizo libremente y no es moralmente responsable de haberlo hecho. Ahora bien, podría parecer que la doctrina de que la coacción y la responsabilidad moral son mutuamente exclusivas no es más que una versión algo particularizada del principio de las posibilidades alternativas. Es natural decir que alguien que ha sido coaccionado a hacer algo no podía hacer otra cosa. Y fácilmente podría parecer que el hecho de ser coaccionado priva a alguien de la libertad y la responsabilidad moral, simplemente porque es un caso especial de imposibilidad de hacer otra cosa. De esta manera, el principio de las posibilidades alternativas puede obtener cierta credibilidad de su asociación con la muy razonable propuesta de que la responsabilidad moral queda excluida por la coacción.

Sin embargo, no está bien que así sea. El hecho de que una persona haya sido coaccionada a actuar como lo hizo puede implicar tanto que no podía hacer otra cosa como que no es responsablemente moral de su acción. Pero su falta de responsabilidad moral no es consecuencia del hecho de no haber podido hacer otra cosa. En otras palabras, no es correcto interpretar la doctrina de que la coacción excluye la responsabilidad moral como una versión particularizada del principio de las posibilidades alternativas.

Supongamos que alguien es amenazado en forma convincente con un castigo que considera inaceptable y que luego hace lo que la persona que lo amenazó le pide. Podemos imaginarnos detalles que nos llevarían a pensar que la persona fue coaccionada para llevar a cabo la acción en cuestión, que no pudo hacer otra cosa y que no es moralmente responsable de haber hecho lo que hizo. No obstante, ¿qué tienen las situaciones de esta clase que garantizan el juicio de que la persona amenazada no es moralmente responsable de su acto?

Esta pregunta puede enfocarse considerando situaciones como la siguiente: Jones tiene razones para decidir hacer algo, luego alguien lo amenaza con un castigo muy severo (tan severo que cualquier persona razonable se rendiría a la amenaza) a menos que haga precisamente aquello, y Jones lo hace. ¿Responsabilizaremos moralmente a Jones por lo que hizo? Creo que eso dependerá de los pa-

peles que creemos que fueron desempeñados, al llevarlo a actuar, por su decisión original y por la amenaza.

Una posibilidad es que $Jones_1$ no sea un hombre razonable; más bien, es un hombre que hace lo que decidió hacer sin importar lo que suceda después y sin importar el costo. En ese caso, la amenaza en realidad no ejerció una fuerza efectiva sobre él. Actuó sin pensar en ella, casi como si no fuera consciente de su existencia. Si, en efecto, así fue, la situación no implicó ninguna coacción. La amenaza no llevó a $Jones_1$ a hacer lo que hizo. Como tampoco fue, de hecho, suficiente para impedir que hiciera otra cosa: si su decisión anterior hubiera sido la de hacer otra cosa, la amenaza no lo habría disuadido en absoluto. Parece evidente que, en estas circunstancias, el hecho de que $Jones_1$ fuera amenazado no reduce de ningún modo la responsabilidad moral que tendría, de otra manera, por su acto. Este ejemplo, sin embargo, no es un contraejemplo de la doctrina que afirma que la coacción libera de responsabilidad ni del principio de las posibilidades alternativas. Pues supusimos que $Jones_1$ es un hombre sobre quien la amenaza no tiene ningún efecto coercitivo y, por ello, que en realidad no lo privó de alternativas a lo que hizo.

Otra posibilidad es que $Jones_2$ fuera empujado a tomar una decisión precipitada por la amenaza. Dada esa amenaza, habría llevado a cabo la acción sin importar qué decisión había tomado. Además, la amenaza lo perturbó tanto que se olvidó por completo de su decisión anterior e hizo todo lo que se le pedía porque tenía terror del castigo con el que había sido amenazado. En este caso, el hecho de que ya hubiera decidido llevarla a cabo no pesa sobre su realización de la acción. Cuando cayó en la cuenta, no pensó en otra cosa que en la amenaza, y sólo el temor lo llevó a actuar. El hecho de que anteriormente $Jones_2$ hubiera decidido, por sus propios motivos, actuar de esa manera puede ser relevante para una evaluación de su carácter; puede tener toda la responsabilidad moral para haber tomado *esa* decisión. Pero no se puede decir que sea moralmente responsable de su acto, puesto que llevó a cabo la acción simplemente como resultado de la coacción a la que fue sometido. Su decisión anterior no desempeñó ningún papel en provocar que hiciera lo que hizo y, por lo tanto, sería gratuito asignarle un papel en la evaluación moral de su acción.

Ahora consideremos una tercera posibilidad. $Jones_3$ no se sintió empujado por la amenaza a tomar una decisión precipitada y tampoco fue indiferente a ella. La amenaza lo impresionó, como impresionaría a cualquier hombre razonable, y él se habría sometido a ella sin reservas si no hubiera tomado ya una decisión que coincidía con la que se le exigía. Sin embargo, de hecho, él llevó a cabo la acción en cuestión sobre la base de la decisión que había tomado antes de recibir la amenaza. Cuando actuó, en realidad no estaba motivado por la amenaza sino únicamente por las consideraciones que lo habían llevado a actuar, aunque sí lo habría estado de no haber tenido suficientes motivos para llevar a cabo la acción en cuestión.

Sin duda, en un caso como éste será muy difícil para cualquiera saber qué pasó exactamente. ¿Acaso $Jones_3$ llevó a cabo la acción en virtud de la amenaza, o sus razones para actuar fueron simplemente las que ya lo habían persuadido de hacerlo? ¿O actuó sobre la base de dos motivos, cada uno de los cuales fue suficiente para su acción? Sin embargo, no es imposible que la situación fuera más clara de lo que son en general las situaciones de esta clase. Y supongamos que para nosotros es obvio que $Jones_3$ actuó sobre la base de su propia decisión y no por la amenaza. Entonces pienso que estaría justificado considerar que su responsabilidad moral por lo que hizo no está afectada por la amenaza aunque, dado que de todas maneras se habría rendido ante la amenaza, no podía haber evitado lo que hizo. Para nosotros sería completamente razonable emitir el mismo juicio con respecto a su responsabilidad que el que habríamos emitido si no hubiéramos tenido conocimiento de la amenaza, puesto que la amenaza de hecho no ejerció influencia alguna sobre su realización de la acción. Él hizo lo que hizo de la misma manera que si no hubiera recibido la amenaza.

III

A primera vista, el caso de $Jones_3$ parecería combinar coacción con responsabilidad moral y, por lo tanto, proporcionar un contraejemplo para la doctrina de que la coacción libera de responsabili-

dad. En realidad, no es tan seguro que lo haga, no obstante, porque no queda claro si el ejemplo constituye una instancia genuina de coacción. ¿Podemos decir que $Jones_3$ fue coaccionado a hacer algo, cuando ya había decidido hacerlo por su cuenta y cuando lo hizo enteramente sobre la base de esa decisión? ¿O sería más correcto decir que $Jones_3$ no fue coaccionado a hacer lo que hizo aunque él mismo reconociera que sentía una fuerza irresistible en virtud de la cual debía hacerlo? Mis propias intuiciones lingüísticas me conducen hacia la segunda alternativa, pero son algo equívocas. Quizá podamos decir cualquiera de estas dos cosas, o quizá debamos añadir una explicación calificativa a cualquiera de las que digamos.

Sin embargo, esta oscuridad no nos impide extraer una importante moraleja de un análisis del ejemplo. Supongamos que decidimos decir que $Jones_3$ *no* fue coaccionado. Nuestra base para decir esto será claramente que es incorrecto considerar que un hombre fue coaccionado a hacer algo a menos que lo haga *a causa de* la fuerza coercitiva ejercida contra él. El hecho de que haya recibido una amenaza irresistible, entonces, no implicará que la persona que la recibe es coaccionada a hacer lo que hace. También será necesario que la amenaza sea la verdadera razón para hacerlo. Por otro lado, supongamos que decidimos decir que $Jones_3$ *fue* coaccionado. Entonces seguramente admitiremos que ser coaccionado no excluye ser moralmente responsable. Y lo más probable es que lleguemos a la perspectiva de que la coacción afecta la responsabilidad moral de una persona sólo cuando la persona actúa como lo hace porque está coaccionada a hacerlo, por ejemplo, cuando el hecho de que ha sido coaccionada da cuenta de su acción.

Sea lo que fuere que decidamos decir, entonces, reconoceremos que la doctrina de que la coacción excluye la responsabilidad moral no es una versión particularizada del principio de las posibilidades alternativas. Las situaciones en las que una persona que hace algo no puede hacer otra cosa porque está sometida a un poder coercitivo o bien no son en absoluto instancias de coacción o bien son situaciones en las que la persona aún puede ser moralmente responsable de lo que hace si no lo hiciera por la coacción. Cuando disculpamos a una persona que ha sido coaccionada, no la disculpamos porque no pudo hacer otra cosa. Aunque alguien sea some-

tido a una fuerza coercitiva que descarta la posibilidad de que lleve a cabo otra acción, puede, a pesar de todo, tener responsabilidad moral por realizar esa acción.

IV

En la medida en que el principio de las posibilidades alternativas deriva su verosimilitud de la asociación con la doctrina de que la coacción excluye la responsabilidad moral, una clara comprensión de esta última disminuye el atractivo de la primera. Sin duda, el caso de $Jones_3$ puede, al parecer, hacer algo más que iluminar la relación entre las dos doctrinas. Quizá también parecería brindar un contraejemplo decisivo del principio de las posibilidades alternativas y, por tanto, mostrar que el principio es falso, puesto que el carácter irresistible de la amenaza a la que $Jones_3$ es sometido bien podría interpretarse en el sentido de que él no puede más que realizar la acción que realiza. Y, sin embargo, la amenaza, ya que $Jones_3$ lleva a cabo la acción sin tenerla en cuenta, no reduce su responsabilidad moral por lo que hace.

Sin duda, se planteará la siguiente objeción contra la sugerencia de que el caso de $Jones_3$ es un contraejemplo del principio de las posibilidades alternativas. Quizás, en cierto sentido, $Jones_3$ no puede realizar otra acción más que la que hace, ya que es un hombre razonable, y la amenaza a la que se enfrenta es suficiente para movilizar a cualquier hombre razonable. Pero no es este sentido el que guarda relación con el principio de las posibilidades alternativas. Su conocimiento de que va a sufrir un castigo intolerable y severo no significa que $Jones_3$, estrictamente hablando, *no pueda* llevar a cabo ninguna otra acción que no sea la que, en efecto, lleva a cabo. Después de todo, aún le queda abierta, y esto es crucial, la posibilidad de desafiar la amenaza si desea hacerlo y aceptar el castigo que su acción le provocaría. En el sentido en que el principio de las posibilidades alternativas emplea el concepto de "podría haber hecho otra cosa", la incapacidad de $Jones_3$ de resistirse a la amenaza no significa que no pueda hacer otra cosa que llevar a cabo la

acción que realiza. De allí que el caso de $Jones_3$ no constituya un ejemplo contrario al principio.

No propongo considerar en qué sentido el concepto de "podría haber hecho otra cosa" figura en el principio de las posibilidades alternativas, como tampoco intentaré medir la fuerza de la objeción que acabo de describir.[1] Porque considero que la fuerza que se piense que pueda tener esta objeción puede ser desviada alterando el ejemplo de la siguiente manera.[2] Supongamos que alguien –digamos Black– quiere que $Jones_4$ lleve a cabo cierta acción. Black está dispuesto a llegar lejos para salirse con la suya, pero prefiere no mostrar sus intenciones innecesariamente. De modo que espera hasta que $Jones_4$ esté a punto de tomar una decisión sobre qué hacer, y no hace nada a menos que le resulte claro (Black es un excelente juez de esas cosas) que $Jones_4$ va a decidirse a hacer algo que *no* sea lo que él quiere que haga. Si resulta claro que $Jones_4$ va a decidirse a hacer otra cosa, Black toma medidas efectivas para asegurarse de que $Jones_4$ se decida a hacer, y que haga, lo que él quiere que haga.[3] Cualesquiera sean las preferencias e inclinaciones iniciales de $Jones_4$, entonces, Black se saldrá con la suya.

1 Los dos conceptos principales empleados en el principio de posibilidades alternativas son "moralmente responsable" y "podría haber hecho otra cosa". Discutir el principio sin analizar alguno de estos conceptos parecería un intento de piratería. El lector debería notar que mi bandera con la calavera y los huesos ha sido desplegada.

2 Después de pensar en el ejemplo que estoy a punto de desarrollar, me enteré de que Robert Nozick, en unas conferencias que dio hace varios años, había formulado un ejemplo del mismo tipo general y lo había propuesto como un contraejemplo del principio de posibilidades alternativas.

3 La suposición de que Black puede predecir lo que $Jones_4$ va a decidir no da por sentada la cuestión del determinismo. Podemos imaginar que $Jones_4$ muchas veces se ha enfrentado a las alternativas A y B a las que ahora se enfrenta, y que su gesto invariablemente se frunció cuando estaba a punto de llevar a cabo A y nunca cuando estaba por decidirse a hacer B. Sabiendo esto y observando el gesto, Black tendría una base para la predicción. Seguramente, esto supone que existe una cierta relación causal entre el estado de $Jones_4$ en el momento del gesto fruncido y sus estados subsiguientes. Sin embargo, cualquier perspectiva razonable de decisión o de acción admitirá que tomar una decisión y llevar a cabo una acción suponen etapas anteriores y posteriores, con relaciones causales entre ellas, y a tal punto que las etapas anteriores no son parte de la decisión ni de la acción. El ejemplo no requiere que estas etapas anteriores estén relacionadas, desde un punto de vista determinista, con sucesos aun anteriores.

¿Qué medidas tomará Black, si cree que debe tomar medidas, a fin de asegurar que $Jones_4$ decida y actúe como él desea? Cualquiera que tenga una teoría con respecto a lo que significa "podría haber hecho otra cosa" está en condiciones de responder a esta pregunta describiendo las medidas que consideraría suficientes para garantizar que, en el sentido pertinente, $Jones_4$ no puede hacer otra cosa. Que Black pronuncie una terrible amenaza y, de esta manera, obligue a $Jones_4$ a llevar a cabo la acción deseada y, a su vez, le impida realizar una prohibida. Que Black le dé a $Jones_4$ una poción o lo hipnotice, y de alguna otra manera similar, genere en $Jones_4$ una compulsión interior irresistible de llevar a cabo el acto que Black desea y de evitar otros. O que Black manipule los procesos mínimos del cerebro y el sistema nervioso de $Jones_4$ de alguna manera más directa, para que las fuerzas causales que entran y salen de sus sinapsis y pasan por los nervios del pobre hombre determinen que elija un acto y que actúe de una manera y no de otra. En otras palabras, cualesquiera sean las condiciones en las que se sostenga que $Jones_4$ no puede hacer otra cosa, que Black haga prevalecer esas condiciones. La estructura del ejemplo es bastante flexible, creo, para hallar una salida a cualquier cargo de irrelevancia acomodando la doctrina sobre la que se basa el cargo.[4]

Ahora supongamos que Black nunca tiene que intervenir porque $Jones_4$, por razones propias, decide llevar a cabo, y lo hace, la acción misma que Black quiere que lleve a cabo. En ese caso, parece claro, $Jones_4$ tendrá precisamente la misma responsabilidad moral por lo que hace que la que habría tenido si Black no hubiera estado dispuesto a tomar medidas para asegurar que lo hiciera. Sería muy poco razonable excusar a $Jones_4$ por su acción, o negar el elogio al que ésta normalmente le daría derecho, sobre la base de que no podía haber hecho otra cosa. Esta situación no lo llevó en absoluto a

4 El ejemplo también es lo suficientemente flexible para permitir la eliminación por completo de Black. Cualquiera que crea que la eficacia del ejemplo se debilita por su dependencia de un manipulador humano que impone su voluntad a $Jones_4$ puede sustituir a Black por una máquina programada para hacer lo que hace Black. Si esto aun no basta, olvidémonos de Black y de la máquina y supongamos que su papel es desempeñado por fuerzas naturales que no involucran ninguna voluntad ni plan en absoluto.

actuar como lo hizo. Habría actuado de la misma manera aunque ello no hubiera ocurrido. En efecto, todo sucedió como habría sucedido sin la presencia de Black en esa situación y sin su disposición para intervenir en ella.

En este ejemplo, hay suficientes condiciones para que $Jones_4$ lleve a cabo la acción en cuestión. La acción que realice no depende de él. Por supuesto, en cierta manera depende de él si actúa por su propia cuenta o como resultado de la intervención de Black. Ello depende de la acción que él mismo esté inclinado a realizar. Pero ya sea que finalmente actúe por su cuenta o como resultado de la intervención de Black, lleva a cabo la misma acción. No tiene más alternativa que hacer lo que Black quiere que haga. Sin embargo, si lo hace por su cuenta, su responsabilidad moral por hacerlo no se ve afectada por el hecho de que Black estuviera acechando con intenciones siniestras, ya que estas intenciones nunca entran en juego.

V

El hecho de que una persona no hubiera podido evitar hacer algo es una condición suficiente para que lo haya hecho. Sin embargo, como muestran algunos de mis ejemplos, es posible que este hecho no desempeñe ningún papel en la explicación de por qué lo hizo. Es posible que no figure en absoluto entre las circunstancias que en realidad provocaron que hiciera lo que hizo, de modo que su acción debe ser explicada sobre otra base por completo diferente. Es decir, aunque la persona hubiera sido incapaz de hacer otra cosa, podría no ser el caso que actuara como lo hizo *porque* no pudo hacer otra cosa. Ahora bien, si alguien no tuvo alternativas para llevar a cabo cierta acción, pero no la realizó porque no pudo hacer otra cosa, entonces habría llevado a cabo exactamente la misma acción aunque *hubiera podido* hacer otra cosa. Las circunstancias que le impidieron hacer otra cosa se podrían haber extraído de la situación sin afectar lo que pasó o por qué pasó en cualquier caso. La razón que efectivamente ha conducido a la persona a hacer lo que hizo o que la obligó

a hacerlo la habría conducido u obligado a hacerlo aun en el caso en que esa persona hubiera podido hacer otra cosa.

Así, no habría habido diferencia, en lo que concierne a su acción o a cómo llegó a realizarla, si las circunstancias que le hicieron imposible evitar realizarla no hubieran prevalecido. Es claro que el hecho de que no haya podido hacer otra cosa no provee ninguna base para suponer que *podría* haber hecho otra cosa de haber sido capaz. Cuando un hecho es, de esta manera, irrelevante respecto del problema de dar cuenta de la acción de una persona, parece infundado asignarle algún peso en la evaluación de su responsabilidad moral. ¿Por qué se debería considerar el hecho al emitir un juicio moral con respecto a la persona cuando no ayuda en absoluto a comprender qué la llevó a actuar como lo hizo o qué podría haber hecho en otras circunstancias?

Ésta, entonces, es la razón por la que el principio de las posibilidades alternativas es erróneo. Afirma que una persona no tiene responsabilidad moral –es decir, que está disculpada– por haber llevado a cabo una acción si hubo circunstancias que le hicieron imposible evitar realizarla. Sin embargo, quizás haya circunstancias que hagan imposible que una persona evite realizar una acción sin que esas circunstancias provoquen en modo alguno que realice esa acción. Ciertamente no sería bueno que la persona se refiriera a circunstancias de esta clase en un intento por absolverse de la responsabilidad moral por llevar a cabo la acción en cuestión. Ello se debe a que, de acuerdo con nuestra hipótesis, en realidad esas circunstancias no tuvieron nada que ver con que la persona haya hecho lo que hizo. Habría hecho exactamente lo mismo y habría sido conducida u obligada a hacerlo de la misma manera, aunque aquéllas no hubieran prevalecido.

Por cierto, con frecuencia disculpamos a las personas por lo que han hecho cuando nos dicen (y les creemos) que no podían hacer otra cosa. Pero esto es porque suponemos que lo que nos dicen sirve para explicar por qué hicieron lo que hicieron. Damos por sentado que no están faltando a la sinceridad, como estaría haciendo alguien que adujera como excusa el hecho de que no tenía modo de evitar hacer lo que hizo, pero que supiera perfectamente que ésta no fue, de ningún modo, la razón por la que lo hizo.

Lo que dije puede sugerir que el principio de las posibilidades alternativas debería revisarse como para afirmar que una persona no es moralmente responsable de lo que ha hecho si lo hizo porque no podía hacer otra cosa. Se observará que esta revisión del principio no afecta seriamente los argumentos de quienes confiaron en el principio original en un intento por sostener que la responsabilidad moral y el determinismo son incompatibles, puesto que si existió determinación causal para que una persona realizara una acción, será verdad que la persona la realizó en virtud de esos determinantes causales. Y si el hecho de que la acción de una persona estuviera determinada causalmente significa que la persona no podía hacer otra cosa, como en general suponen los filósofos que defienden la tesis de la incompatibilidad, entonces el hecho de que la acción de una persona estuviera determinada causalmente significará que la persona la realizó porque no podía hacer otra cosa. El principio revisado de las posibilidades alternativas implicará, en cuanto a esta suposición con respecto al significado de "podía haber hecho otra cosa", que una persona no es moralmente responsable de lo que hizo si su acción estuvo determinada causalmente. No obstante, no creo que esta revisión del principio sea aceptable.

Supongamos que una persona nos dice que hizo lo que hizo porque no podía hacer otra cosa; o supongamos que hace una afirmación similar: que hizo lo que hizo porque tenía que hacerlo. Con frecuencia aceptamos enunciados como ésos (si los creemos) como excusas válidas, y esos enunciados, a primera vista, parecen invocar el principio revisado de las posibilidades alternativas. Pero creo que cuando aceptamos enunciados de ese tipo como excusas válidas es porque suponemos que nos dicen más de lo que los enunciados transmiten estricta y literalmente. Entendemos que la persona que ofrece la excusa quiere decir que hizo lo que hizo *sólo porque* no podía hacer otra cosa o *sólo porque* tenía que hacerlo. Y entendemos que quiere decir, más en particular, que cuando hizo lo que hizo no fue porque eso fuera lo que realmente quería hacer. El principio de las posibilidades alternativas, entonces, debería ser reemplazado, en mi opinión, por el siguiente principio: una persona no es moralmente responsable de lo que hizo si lo hizo sólo

porque no podía hacer otra cosa. Este principio no parece estar en conflicto con la visión de que la responsabilidad moral es compatible con el determinismo.

Lo siguiente puede ser verdad: existieron circunstancias que hicieron imposible que una persona evitara hacer algo; estas circunstancias, en realidad, desempeñaron un papel para provocar que esa persona hiciera lo que hizo, de modo que es correcto decir que lo hizo porque no podía hacer otra cosa; la persona realmente quería hacer lo que hizo; lo hizo porque era lo que realmente quería hacer, de modo que no es correcto decir que hizo lo que hizo sólo porque no podía hacer otra cosa. En estas condiciones, la persona bien puede ser moralmente responsable de lo que hizo. Por otro lado, no será moralmente responsable de lo que hizo si lo hizo sólo porque no podía hacer otra cosa, aunque lo que hizo fuera algo que realmente quería hacer.

2
La libertad de la voluntad y el concepto de persona

Lo que los filósofos han aceptado últimamente como análisis del concepto de persona no es, en absoluto, el análisis de *ese* concepto. Strawson, cuya manera de emplearlo representa la pauta generalizada en la actualidad, identifica el concepto de persona con "el concepto de un tipo de entidad tal que *tanto* los predicados que le atribuyen estados de conciencia *como* los predicados que le atribuyen características corpóreas [...] pueden ser igualmente aplicados a un individuo único de ese tipo único".[1] Sin embargo, hay muchas entidades aparte de las personas que tienen propiedades tanto mentales como físicas. Da la casualidad –aunque parece extraordinario que así sea– de que no hay una palabra común en inglés para el tipo de entidad que Strawson tiene en mente, un tipo que incluye no sólo a seres humanos, sino también a animales de diversas especies menores. No obstante, ello no justifica el mal uso de un término filosófico valioso.

La cuestión de si los miembros de algunas especies animales son personas no va a ser resuelta, seguramente, mediante la mera determinación de si es correcto emplear para ellos, además de

1 P. F. Strawson, *Individuals*, Londres, Methuen, 1959, pp. 101-102 [trad. esp.: *Individuos*, Madrid, Taurus, 1989.] El uso que Ayer hace de "persona" es similar: "Es característico de las personas, en este sentido, que aparte de tener diversas propiedades físicas [...] también se les reconozcan diversas formas de conciencia", en A. J. Ayer, *The Concept of a Person*, Nueva York, St. Martin's, 1963, p. 82 [trad. esp.: *El concepto de persona*, Barcelona, Seix Barral, 1966]. Lo que les preocupa a Strawson y a Ayer es el problema de comprender la relación entre la mente y el cuerpo, y no el problema, bastante diferente, de comprender qué significa ser una criatura que no sólo tiene una mente y un cuerpo, sino que también es una persona.

predicados que les atribuyan características corpóreas, predicados que les atribuyan estados de conciencia. De hecho, atenta contra nuestro lenguaje emplear el término "persona" para aquellas numerosas criaturas que tienen tanto propiedades psicológicas como materiales, pero que, evidentemente, no son personas en ningún sentido normal de la palabra. Este mal uso del lenguaje es, sin duda, ajeno a cualquier error teórico. Sin embargo, aunque la ofensa es "meramente verbal", causa un daño importante, debido a que reduce en forma gratuita nuestro vocabulario filosófico y aumenta la probabilidad de que pasemos por alto la importante área de investigación con la que más naturalmente se asocia el término "persona". Se podría haber esperado que para los filósofos ningún problema presentara un mayor y más constante interés que el de comprender lo que somos en esencia. No obstante, por lo general se descuida tanto este problema que es habitual llevarse el nombre por delante casi sin que se note y, evidentemente, sin provocar ningún sentimiento generalizado de pérdida.

Hay un sentido en que la palabra "persona" es simplemente la forma singular de "gente" y en que ambos términos sólo connotan la pertenencia a cierta especie biológica. Sin embargo, en los sentidos de la palabra que son de mayor interés filosófico, los criterios para ser una persona no sirven fundamentalmente para distinguir a los miembros de nuestra propia especie de los miembros de otras especies. Por el contrario, están formulados para captar aquellos atributos que son el centro de nuestra preocupación más humana con nosotros mismos y la fuente de lo que consideramos más importante y más problemático en nuestras vidas. Ahora bien, estos atributos tendrían igual importancia para nosotros incluso si no fueran, de hecho, peculiares y comunes a los miembros de nuestra propia especie. Lo que más nos interesa de la condición humana no nos interesaría menos si también fuera un rasgo característico de la condición de otras criaturas.

Nuestro concepto de nosotros como personas no se puede comprender, por tanto, como un concepto de atributos que necesariamente son específicos de la especie. Es posible, desde el punto de vista conceptual, que miembros de especies no humanas nuevas –o

incluso conocidas– sean personas; y también es posible, desde el mismo punto de vista, que algunos miembros de la especie humana no sean personas. En realidad suponemos, por otra parte, que ningún miembro de otra especie es una persona. De acuerdo con esto, se presume que las personas se definen esencialmente por una serie de características que –ya sea correcta o incorrectamente– en general suponemos específicamente humanas.

Creo que una diferencia esencial entre las personas y otras criaturas puede encontrarse en la estructura de la voluntad de una persona. Los seres humanos no son los únicos que tienen deseos y motivaciones ni los únicos que pueden elegir. Comparten estas cosas con los miembros de algunas otras especies, algunos de los cuales incluso parecen deliberar y tomar decisiones basadas en un pensamiento previo. No obstante, parece ser peculiarmente característico de los seres humanos el que puedan formar lo que denominaré "deseos de segundo orden" o "deseos del segundo orden".

Además de querer, elegir y ser inducidos a *hacer* esto o aquello, es posible que los hombres también quieran tener (o no) ciertos deseos y motivaciones. Son capaces de querer ser diferentes, en sus preferencias y en sus propósitos, de lo que son. Muchos animales parecen tener la capacidad de lo que denominaré "deseos de primer orden" o "deseos del primer orden", que simplemente son deseos de hacer o no una cosa u otra. Sin embargo, ningún animal, salvo el hombre, parece tener la capacidad de realizar la autoevaluación reflexiva que se manifiesta en la formación de los deseos de segundo orden.[2]

2 En pos de la simplicidad, sólo me referiré a lo que alguien quiere o desea, sin tener en cuenta los fenómenos relacionados, como las elecciones y las decisiones. Propongo usar los verbos "querer" (*to want*) y "desear" (*to desire*) en forma indistinta, aunque de ninguna manera sean sinónimos perfectos. La razón de mi renuncia a los matices establecidos de estas palabras surge del hecho de que el verbo "querer", que es más apropiado para mi objetivo en cuanto a su significado, no se presta con tanta facilidad a la formación de sustantivos como sucede con el verbo "desear". En inglés, quizá sea aceptable, aunque poco elegante, hablar en plural de *the wants* (los "quereres") de alguien. Sin embargo, hablar en singular de *the want* (el "querer") de una persona resultaría una abominación.

I

El concepto que designa el verbo "querer" es en extremo esquivo. Una afirmación tal como "*A* quiere hacer *X*" –tomada en forma aislada, separada de un contexto que sirva para ampliar o especificar su significado– transmite extraordinariamente escasa información. Una afirmación como ésa puede ser coherente, por ejemplo, con cada una de las siguientes: a) la posibilidad de hacer *X* no provoca en *A* una sensación o respuesta emocional susceptible de introspección; b) *A* no es consciente de que quiere hacer *X*; c) *A* cree que no quiere hacer *X*; d) *A* quiere abstenerse de hacer *X*; e) *A* quiere hacer *Y* y cree que es imposible que pueda hacer tanto *Y* como *X*; f) *A* "en realidad" no quiere hacer *X*; g) *A* preferiría morir antes que hacer *X*; y así sucesivamente. Por tanto, no basta con formular la distinción entre los deseos de primer orden y los de segundo orden, como he hecho yo, sugiriendo simplemente que alguien tiene un deseo de primer orden cuando quiere hacer o no hacer tal cosa y que tiene un deseo de segundo orden cuando quiere tener o no tener cierto deseo del primer orden.

Según las interpretaré, afirmaciones tales como "*A* quiere hacer *X*" abarcan una gama de posibilidades bastante amplia.[3] Es posible que sean verdaderas incluso cuando son verdaderas afirmaciones como las identificadas con las letras (a) a (g): cuando *A* no es consciente de ningún sentimiento respecto de hacer *X*, cuando no sabe que quiere hacer *X*, cuando se engaña acerca de lo que quiere y cree falsamente que no quiere hacer *X*, cuando también tiene otros deseos que están en conflicto con su deseo de hacer *X*, o cuando es ambivalente. Los deseos en cuestión pueden ser conscientes o inconscientes, no es necesario que sean unívocos, y *A* puede estar errado acerca de ellos. Sin embargo, existe otra fuente de incertidumbre respecto de afirmaciones que identifican los de-

3 Lo que expongo en este párrafo se aplica no sólo a los casos en que "hacer *X*" se refiere a una acción o a una inacción posibles. También se aplica a los casos en que "hacer *X*" se refiere a un deseo de primer orden y en los que la afirmación "*A* quiere hacer *X*" es, por tanto, una versión abreviada de una afirmación –"*A* quiere querer hacer *X*" – que identifica un deseo del segundo orden.

seos de alguien y, en este punto, es importante para mi propósito que sea menos permisivo.

Consideremos, en primera instancia, aquellas afirmaciones de la forma "*A* quiere hacer *X*" que identifican deseos de primer orden, es decir, afirmaciones en las que el término "hacer *X*" se refiere a una acción. Una afirmación de este tipo no indica, por sí misma, la fuerza relativa del deseo de *A* de hacer *X*. No deja en claro si es posible que este deseo desempeñe un papel decisivo en lo que *A* en realidad hace o intenta hacer. Por ello, se podría decir acertadamente que *A* quiere hacer *X* aun cuando su deseo de hacer *X* sea sólo uno entre sus deseos y aun cuando éste diste mucho de ser el principal. Por consiguiente, puede ser verdad que *A* quiera hacer *X* aunque en realidad prefiera hacer otra cosa; y puede ser verdad que quiera hacer *X* a pesar de que, cuando actúe, no sea el deseo de hacer *X* lo que lo motiva a hacer lo que hace. Por otro lado, alguien que afirma que *A* quiere hacer *X* podría querer expresar que es este deseo el que está motivando o induciendo a *A* a hacer lo que realmente está haciendo o que *A*, de hecho, será inducido por este deseo (a menos que cambie de parecer) cuando actúe.

Dado el uso especial de *voluntad* que propongo adoptar, sólo cuando la afirmación se emplea de la segunda manera identifica la voluntad de *A*. Identificar la voluntad de un agente significa ya identificar el deseo (o los deseos) que lo inducen a realizar alguna acción que lleva a cabo, ya identificar el deseo (o los deseos) que lo inducirán o lo inducirían cuando actúe o si actúa. Por tanto, la voluntad de un agente coincide con uno o más de sus deseos de primer orden. Pero la noción de voluntad, según la estoy empleando, no es coextensiva con la noción de deseos de primer orden. No es la noción de algo que simplemente inclina a un agente, en cierto grado, a actuar de cierta manera. Es, más bien, la noción de un deseo *efectivo*, tal que que induce (o inducirá o induciría) a una persona a llevar a cabo la acción. Por tanto, la noción de voluntad no es coextensiva con la noción de lo que un agente tiene intenciones de hacer, ya que, aunque alguien pueda tener una intención firme de hacer *X*, tiene la posibilidad de hacer otra cosa en vez de hacer *X* porque, a pesar de su intención, su deseo de hacer *X* resulta ser más débil o menos efectivo que otro en conflicto con el primero.

Ahora consideremos las afirmaciones de la forma "*A* quiere hacer *X*" que identifican deseos de segundo orden, es decir, afirmaciones en las que el término "hacer *X*" se refiere a un deseo del primer orden. Hay también dos tipos de situaciones en que podría ser verdad que *A* quiere querer hacer *X*. En primer lugar, podría ser verdad que *A* quiere tener el deseo de hacer *X*, a pesar de tener un deseo unívoco, sin ningún tipo de conflicto ni ambivalencia, de abstenerse de hacer *X*. En otras palabras, alguien podría tener cierto deseo y, sin embargo, querer unívocamente que ese deseo no se satisfaga.

Supongamos que un médico especialista en psicoterapia con drogadictos cree que su capacidad de ayudar a sus pacientes mejoraría si pudiera comprender mejor qué significa para ellos desear la droga a la que son adictos. Supongamos que esto lo lleva a querer desear la droga. Si lo que quiere es sentir un deseo genuino, no sólo se trata, entonces, de experimentar las sensaciones que, por lo general, tienen los adictos cuando son dominados por sus deseos de consumir la droga. Lo que el médico quiere, en la medida en que quiere tener el deseo, es verse inclinado o inducido, en cierta medida, a consumirla.

Sin embargo, es totalmente posible que, aunque quiera sentirse inducido por el deseo de consumir la droga, no quiera que su deseo sea efectivo. Es posible que no quiera que el deseo lo induzca a la acción. No es necesario que esté interesado en descubrir qué se siente al consumir la droga. Y, en la medida en que ahora sólo quiere *querer* consumirla y no *consumirla*, no hay nada en lo que quiere en este momento que pueda ser satisfecho con la droga en sí misma. De hecho, es posible que ahora tenga un deseo totalmente unívoco de *no* consumir la droga; y es posible que disponga, prudentemente, que le sea imposible satisfacer el deseo que tendría si su deseo de querer la droga fuera satisfecho con el tiempo.

Por tanto, sería incorrecto inferir, a partir del hecho de que el médico ahora quiere desear consumir la droga, que ya tiene deseos de consumirla. Su deseo –de segundo orden– de ser inducido a consumir la droga no implica que sienta el deseo –de primer orden– de consumirla. Si se le fuera a administrar la droga, ello podría no satisfacer ningún deseo implícito en su deseo de querer

consumirla. Si bien quiere querer consumir la droga, es posible que *no* desee consumirla; podría ser que *todo* lo que quiere es probar el deseo de consumirla. Es decir, su deseo de tener cierto deseo que no tiene puede no constituir un deseo de que su voluntad sea en absoluto diferente de lo que es.

Alguien que, sólo de esta forma trunca, quiere querer hacer *X* se ubica en el margen del preciosismo, y el hecho de que quiere querer hacer *X* no es pertinente para la identificación de su voluntad. Sin embargo, hay un segundo tipo de situación que puede describirse como "*A* quiere querer hacer *X*"; y cuando se emplea la afirmación para describir una situación de este segundo tipo, en este caso sí se refiere a lo que *A* quiere que sea su voluntad. En dichos casos, la afirmación significa que *A* quiere que el deseo de hacer *X* sea el deseo que lo induzca a actuar efectivamente. No es sólo que quiere que el deseo de hacer *X* se encuentre entre los deseos por los cuales, en una u otra medida, es inducido o se siente inclinado a actuar. Quiere que este deseo sea efectivo, es decir, que proporcione la motivación de lo que realmente hace. Ahora bien, cuando la afirmación de que *A* quiere querer hacer *X* se emplea de esta manera, implica que *A* ya tiene el deseo de hacer *X*. No podría ser verdad a la vez que *A* quiere que el deseo de hacer *X* lo induzca a la acción y que no quiere hacer *X*. Sólo si realmente quiere hacer *X* podrá querer en forma coherente que el deseo de hacer *X* no sea sólo uno de sus deseos, sino que sea más decididamente su voluntad.[4]

Supongamos que un hombre quiere ser motivado en lo que hace por el deseo de concentrarse en su trabajo. Si esta suposición es correcta, necesariamente es verdad que ya quiere concentrarse en su

4 No está tan claro que la relación de implicación aquí descripta se mantenga en ciertos tipos de casos que, pienso, podrían ser considerados con imparcialidad no estándar, donde la diferencia esencial entre los casos estándar y los no estándar reside en el tipo de descripción mediante la cual se identifica el deseo de primer orden en cuestión. Así, supongamos que *A* admira a *B* en forma tan exagerada que, incluso aunque no sepa qué quiere hacer *B*, *A* quiere ser inducido en forma efectiva por cualquier deseo que induzca en forma efectiva a *B;* en otras palabras, sin saber cuál es la voluntad de *B*, *A* quiere que su propia voluntad sea la misma. Por supuesto, no se desprende de ello que *A* ya tenga, entre sus deseos, un deseo como el que constituye la voluntad de *B*. No proseguiré aquí con preguntas acerca de si existen contraejemplos genuinos para la afirmación hecha en el texto o si, en caso de que existieran, esa afirmación debería modificarse.

trabajo. Este deseo es nuevo entre sus deseos. Sin embargo, la cuestión de si se cumple o no su deseo de segundo orden no depende sólo de si el deseo que quiere es uno de sus deseos. Depende de si este deseo es, tal como él quiere que sea, su deseo efectivo o voluntad. Si, a la hora de la verdad, es su deseo de concentrarse en su trabajo lo que lo induce a hacer lo que hace, entonces lo que quiere en ese momento es sin duda (en el sentido pertinente) lo que quiere querer. Por otra parte, si el deseo que lo induce realmente cuando actúa es algún otro, entonces lo que quiere en ese momento no es (en el sentido pertinente) lo que quiere querer. Ello será así aunque el deseo de concentrarse en su trabajo continúe estando entre sus deseos.

II

Alguien tiene un deseo del segundo orden ya sea cuando simplemente quiere tener cierto deseo, ya cuando quiere que cierto deseo sea su voluntad. En situaciones de este último tipo, denominaré sus deseos de segundo orden "voliciones de segundo orden" o "voliciones del segundo orden". Ahora bien, considero que es esencial tener voliciones de segundo orden y no deseos de segundo orden en general para ser una persona. Desde el punto de vista lógico es posible, aunque improbable, que exista un agente con deseos de segundo orden, pero sin voliciones del segundo orden. Dicha criatura, en mi opinión, no sería una persona. Emplearé el término *inconsciente** para referirme a agentes que tienen deseos de primer orden, pero que no son personas, porque, tengan o no deseos del segundo orden, no tienen voliciones de segundo orden.[5]

* Del inglés *wanton*. En español, no se cuenta con un equivalente exacto de *wanton*, ya que en inglés esta palabra también incluye los distintos matices de "licencioso", "libertino", "desenfrenado", "displicente", "ndisciplinado"e, incluso, "caprichoso". [N. de T.]

5 Las criaturas con deseos de segundo orden, pero sin voliciones de segundo orden difieren en gran medida de los animales brutos y, para algunos fines, sería conveniente considerarlas personas. Mi propuesta, que rechaza la denominación "persona" para ellos, es, por tanto, algo arbitraria. En gran parte, la adopto porque

La característica esencial de un agente inconsciente es que no le importa su voluntad. Sus deseos lo inducen a hacer ciertas cosas, sin que sea verdad ni que quiere ser inducido por esos deseos ni que prefiere ser inducido por otros deseos. La clase de agentes inconscientes incluye a todos los animales no humanos que tienen deseos y a todos los niños muy pequeños. Quizá también incluya a algunos seres humanos adultos. En todo caso, los seres humanos adultos pueden ser más o menos inconscientes; es posible que actúen de manera inconsciente en respuesta a deseos de primer orden respecto de los cuales no tienen voliciones del segundo orden, con mayor o menor frecuencia.

El hecho de que un agente inconsciente no tenga voliciones de segundo orden no significa que cada uno de sus deseos de primer orden se traduzca en forma irreflexiva y de inmediato en acción. Quizá no tenga la oportunidad de actuar de acuerdo con algunos de sus deseos. Más aun, la traducción de sus deseos en acción puede ser demorada o impedida ya sea por deseos conflictivos del primer orden, ya por la deliberación. Esto se debe a que es posible que un agente inconsciente posea y emplee facultades racionales de una jerarquía alta. No hay nada en el concepto de agente inconsciente que implique que éste no pueda razonar o deliberar acerca de cómo hacer lo que quiere hacer. Lo que distingue al agente inconsciente racional de otros agentes racionales es que no le preocupa la conveniencia de sus deseos. Hace caso omiso de la pregunta acerca de cuál ha de ser su voluntad. No sólo sigue cualquier proceder que se vea más fuertemente inclinado a seguir, sino que no le importa cuál de sus inclinaciones es la más fuerte.

Por tanto, una criatura racional, que reflexiona acerca de la conveniencia de sus deseos de seguir un curso de acción u otro, puede ser, sin embargo, inconsciente. Al sostener que la esencia de ser una persona reside no en la razón sino en la voluntad, estoy lejos de sugerir que una criatura sin razón puede ser una persona. Ello se

---

facilita la formulación de algunas de las observaciones que quiero hacer. De aquí en adelante, cada vez que considere afirmaciones tales como "*A* quiere querer hacer *X*", tendré en cuenta afirmaciones que identifiquen voliciones de segundo orden y no afirmaciones que identifiquen deseos de segundo orden que no son voliciones de segundo orden.

debe a que sólo en virtud de sus capacidades racionales una persona es capaz de volverse críticamente consciente de su propia voluntad y de formar voliciones del segundo orden. La estructura de la voluntad de una persona presupone, en consecuencia, que es un ser racional.

La distinción entre una persona y un agente inconsciente puede ilustrarse mediante la diferencia entre dos drogadictos. Supongamos que la condición fisiológica responsable de la adicción es la misma en ambos hombres, y que ambos sucumben de manera inevitable a sus deseos periódicos de consumir la droga a la cual son adictos. Uno de los adictos detesta su adicción y siempre lucha desesperadamente, aunque en vano, contra su embestida. Prueba todo lo que piensa que podría permitirle superar sus deseos de consumir la droga. Sin embargo, estos deseos son demasiado poderosos para él y, al final, invariablemente lo conquistan. Se trata de un adicto contra su voluntad, indefenso ante sus propios deseos.

El adicto contra su voluntad tiene deseos conflictivos de primer orden: quiere consumir la droga y también quiere abstenerse de hacerlo. Sin embargo, además de estos deseos de primer orden, tiene una volición del segundo orden. No es neutral respecto del conflicto entre su deseo de consumir la droga y su deseo de abstenerse de hacerlo. Este último deseo, y no el primero, es el que él quiere que constituya su voluntad; el último deseo y no el primero es el que él quiere que sea efectivo y le brinde el propósito que intentará concretar en lo que haga realmente.

El otro adicto es un agente inconsciente. Sus acciones reflejan la economía de sus deseos de primer orden, sin que le preocupe si los deseos que lo inducen a actuar son los deseos por los cuales quiere ser inducido a actuar. Si tiene problemas para obtener la droga o para administrársela, sus respuestas a su necesidad de consumirla podrían requerir deliberación. Sin embargo, nunca se le ocurre considerar si quiere que las relaciones entre sus deseos tengan como resultado que él tenga la voluntad que tiene. El adicto inconsciente puede ser un animal y, por tanto, ser incapaz de preocuparse por su voluntad. En cualquier caso, en lo que concierne a la inconciencia de su falta de preocupación, no es diferente de un animal.

El segundo de estos adictos puede sufrir un conflicto de primer orden similar al conflicto del primer orden que sufre el primer adicto. Sea humano o no, es posible que el agente inconsciente (quizá debido al condicionamiento) quiera tanto consumir la droga como abstenerse de consumirla. A diferencia del adicto contra su voluntad, sin embargo, no prefiere que uno de sus deseos conflictivos prevalezca sobre el otro; no prefiere que un deseo de primer orden, y no otro, constituya su voluntad. Sería engañoso decir que es neutral en cuanto al conflicto entre sus deseos, puesto que ello sugeriría que los considera igualmente aceptables. Y ya que no tiene otra identidad que sus deseos de primer orden, no es cierto ni que prefiera uno antes que el otro ni que prefiera no tomar partido.

Para el adicto contra su voluntad, que es una persona, importa mucho cuál de sus deseos conflictivos de primer orden gana. Ambos deseos son suyos, por cierto; y si finalmente consume la droga o finalmente logra abstenerse de consumirla, actúa para satisfacer lo que es, en sentido literal, su propio deseo. En cualquiera de los casos, hace algo que él mismo quiere hacer, y no lo hace debido a cierta influencia externa cuyo objetivo coincide por casualidad con el propio, sino por su deseo de hacerlo. No obstante, el adicto contra su voluntad se identifica a sí mismo mediante la formación de una volición de segundo orden con uno de sus deseos conflictivos de primer orden y no con el otro. Hace que uno de ellos sea más auténticamente suyo y, al hacerlo, se distancia del otro. En virtud de esta identificación y este distanciamiento, que se logran mediante la formación de una volición de segundo orden, el adicto contra su voluntad puede, en forma significativa, hacer las declaraciones –desconcertantes desde el punto de vista analítico– de que la fuerza que lo induce a consumir la droga es una fuerza distinta de la propia, y que esta fuerza lo induce a consumirla no por su propio libre albedrío, sino, por el contrario, contra su voluntad.

Al adicto inconsciente no le puede importar o no le importa cuál de sus deseos conflictivos de primer orden gana. Su falta de preocupación no se debe a su incapacidad de encontrar una base convincente para sus preferencias. Se debe o bien a su falta de capacidad de reflexión o bien a su indiferencia mecánica frente a la

tarea de evaluar sus propios deseos y motivos.[6] Hay una sola cuestión en la lucha a la cual puede conducir este conflicto de primer orden: cuál de sus deseos conflictivos es más fuerte. Debido a que es inducido por ambos deseos, no sentirá plena satisfacción por lo que haga independientemente de cuál de ellos sea efectivo. Pero *para él* es lo mismo si se imponen sus anhelos o su aversión. A él no le interesa el conflicto que existe entre ellos y, por tanto, a diferencia del adicto contra su voluntad, no puede ni ganar ni perder la lucha en la que está empeñado. Cuando una *persona* actúa, el deseo por el cual es inducido a la acción es o bien la voluntad que quiere o bien una voluntad que no quiere tener. Cuando un *agente inconsciente* actúa, no se trata de ninguna de las dos.

III

Existe una relación muy estrecha entre la capacidad de formar voliciones de segundo orden y otra capacidad que es esencial para las personas: una que a menudo ha sido considerada una marca distintiva de la condición humana. El hecho de que una persona sea capaz tanto de disfrutar como de carecer de la libertad de la voluntad se debe sólo a que tiene voliciones del segundo orden. El concepto de persona, por consiguiente, no es sólo el concepto de un tipo de entidad que tiene tanto deseos de primer orden como voliciones del segundo orden. También se lo puede interpretar como el concepto de un tipo de entidad para la que la libertad de su voluntad podría representar un problema. Este concepto ex-

6 Cuando digo que la evaluación de sus propios deseos y de sus motivos es característica de una persona, no quiero sugerir que las voliciones de segundo orden de una persona manifiesten necesariamente una postura *moral* de su parte respecto de sus deseos de primer orden. Es posible que la persona no evalúe sus deseos de primer orden desde el punto de vista de la moralidad. Más aun, es posible que una persona actúe en forma caprichosa e irresponsable al formar sus voliciones de segundo orden y no considere con seriedad qué está en juego. Las voliciones de segundo orden expresan evaluaciones sólo en el sentido de que son preferencias. No hay una restricción esencial en cuanto al tipo de fundamento, si es que lo hay, sobre el cual se forman.

cluye a todos los agentes inconscientes, tanto infrahumanos como humanos, puesto que no satisfacen una condición esencial para gozar de la libertad de la voluntad. Y excluye a esos seres sobrehumanos, si los hay, cuya voluntad es necesariamente libre.

De hecho, ¿qué tipo de libertad es la libertad de la voluntad? Esta pregunta requiere una identificación del área especial de la experiencia humana con la que el concepto de libertad de la voluntad, a diferencia de los conceptos de otros tipos de libertad, está particularmente vinculado. Al abordarla, mi objetivo será, en primer lugar, localizar el problema con el que se enfrenta una persona de manera inmediata cuando se interesa por la libertad de su voluntad.

Según una tradición filosófica conocida, ser libre es, fundamentalmente, cuestión de hacer lo que uno quiere hacer. Ahora bien, la noción de un agente que hace lo que quiere hacer no es, de ninguna manera, totalmente clara: tanto el hacer como el querer, como también la adecuada relación entre ellos, tienen que dilucidarse. Pero a pesar de que esta noción requiere una mayor exactitud y su formulación tiene que refinarse, creo que al menos captura parte de lo que está implícito en la idea de un agente que *actúa* libremente. No obstante, pierde por completo el contenido peculiar de la idea –bastante diferente– de un agente cuya *voluntad* es libre.

No suponemos que los animales gozan de libertad de la voluntad, aunque reconocemos que un animal puede ser libre para correr en la dirección que quiera. Por consiguiente, tener la libertad de hacer lo que uno quiere hacer no es una condición suficiente para tener libre albedrío. Tampoco es una condición necesaria, puesto que privar a alguien de su libertad de acción no significa necesariamente debilitar la libertad de su voluntad. Cuando un agente es consciente de que hay ciertas cosas que no es libre de hacer, ello, sin dudas, afecta sus deseos y limita el rango de las elecciones que puede hacer. Pero supongamos que alguien, sin saberlo, ha perdido o ha sido privado de su libertad de acción. Aunque ya no tiene la libertad de hacer lo que quiere hacer, es probable que su voluntad siga siendo tan libre como lo era anteriormente. A pesar del hecho de que no es libre para traducir sus deseos en acciones o para actuar según lo que determina su voluntad, aún podrá formar

esos deseos y tomar esas determinaciones tan libremente como si su libertad de acción no hubiera sido afectada.

Cuando preguntamos si la voluntad de una persona es libre, no estamos preguntando si tiene la posibilidad de traducir sus deseos de primer orden en acciones. Eso sería preguntar si es libre para hacer lo que le plazca. La pregunta por la libertad de su voluntad no tiene que ver con la relación entre lo que hace y lo que quiere hacer. Más bien, tiene que ver con sus deseos en sí mismos. Pero, ¿qué sucede con ellos?

Me parece natural y útil interpretar la pregunta de si la voluntad de una persona es libre en estrecha analogía con la pregunta de si un agente goza de libertad de acción. Ahora bien, la libertad de acción es (aproximadamente, al menos) la libertad de hacer lo que uno quiere hacer. De manera similar, entonces, la afirmación de que una persona goza de libertad de la voluntad significa (también aproximadamente) que es libre de querer lo que quiere querer. Para ser más exacto, significa que es libre de desear lo que quiera desear, o de tener el deseo que quiera. Así como la cuestión de la libertad de acción de un agente está relacionada con el hecho de si se trata de la acción que quiere realizar, la cuestión de la libertad de su voluntad está relacionada con el hecho de si es la voluntad que quiere tener.

Por tanto, una persona ejercita la libertad de la voluntad asegurando la conformidad de su voluntad con sus voliciones de segundo orden. Y la discrepancia entre su voluntad y sus voliciones de segundo orden o el hecho de saber que su coincidencia no es obra suya, sino sólo una casualidad, hacen que la persona que no tiene esta libertad sienta su carencia. La voluntad del adicto contra su voluntad no es libre. Esto se ve en el hecho de que no es la voluntad que quiere. También es verdad, aunque de manera diferente, que la voluntad del adicto inconsciente no es libre. El adicto inconsciente no tiene ni la voluntad que quiere ni una voluntad diferente de la voluntad que quiere. Dado que no tiene voliciones del segundo orden, la libertad de su voluntad no puede ser un problema para él. Desde siempre, por así decirlo, carece de ella.

Por lo general, la gente es mucho más complicada de lo que podría sugerir mi esbozo sobre la estructura de la voluntad de una

persona. Hay tantas oportunidades para las ambivalencias, el conflicto y el autoengaño respecto de los deseos del segundo orden, por ejemplo, como respecto de los deseos de primer orden. Si hay un conflicto no resuelto entre los deseos de segundo orden de alguien, esta persona está en peligro de no tener volición de segundo orden, puesto que, a menos que se resuelva este conflicto, no tiene preferencias respecto de cuál de sus deseos de primer orden ha de ser su voluntad. Esta condición, si es tan seria como para impedirle identificarse, en forma lo suficientemente decisiva, con *alguno* de sus deseos conflictivos de primer orden, lo destruye como persona. Esto se debe a que tiende a paralizar su voluntad y a impedir que actúe o tiende a distanciarlo de su voluntad de manera que ésta opera sin su participación. En ambos casos, tal como el adicto contra su voluntad aunque en forma diferente, se convierte en un espectador indefenso de las fuerzas que lo inducen.

Otra cuestión compleja es que la persona puede tener, en especial si sus deseos de segundo orden están en conflicto, deseos y voliciones de un orden superior al segundo. No existe un límite teórico para la extensión de la serie de deseos de órdenes más y más altos; no existe nada, salvo el sentido común y, quizá, una fatiga salvadora que impide que un individuo se niegue, en forma obsesiva, a identificarse con alguno de sus deseos hasta que forma un deseo del orden superior siguiente. La tendencia a generar una serie tal de actos de formación de deseos, que constituiría un caso de humanización sin freno, también conduce a la destrucción de una persona.

No obstante, es posible poner fin a tal serie de actos sin cortarla en forma arbitraria. Cuando una persona se identifica *decididamente* con uno de sus deseos de primer orden, este compromiso "resuena" a través de la serie potencialmente infinita de órdenes superiores. Consideremos una persona que, sin reserva ni conflictos, quiere ser motivada por el deseo de concentrarse en su trabajo. El hecho de que su volición de segundo orden de ser inducida por este deseo es decisiva significa que no hay lugar para preguntas acerca de la pertinencia de deseos o voliciones de órdenes superiores. Supongamos que se le pregunta a la persona si quiere querer querer concentrarse en su trabajo. Ella bien puede insistir en

que no se plantea la pregunta sobre un deseo de tercer orden. Sería erróneo sostener que, debido a que no ha considerado si quiere la volición de segundo orden que ha formado, es indiferente a la pregunta de si quiere que su voluntad concuerde con esta volición o con alguna otra. La firmeza del compromiso que ha contraído significa que ha decidido que no queda por formular ninguna otra pregunta acerca de su volición de segundo orden, en ningún orden superior. Es relativamente poco importante la cuestión de si lo explicamos diciendo que este compromiso genera, en forma implícita, una serie infinita de deseos de confirmación de órdenes superiores o que el compromiso es equivalente a una disolución de la importancia de todas las preguntas relacionadas con órdenes superiores de deseos.

Ejemplos como el del adicto contra su voluntad podrían sugerir que las voliciones del segundo orden, o de órdenes superiores, deben formarse de manera deliberada y que, característicamente, una persona lucha para asegurar que se satisfagan. Pero la conformidad de la voluntad de una persona a sus voliciones de orden superior puede ser mucho más irreflexiva y espontánea. Algunas personas son inducidas en forma natural por la bondad cuando quieren ser buenas, y por la maldad cuando quieren ser malas, sin ninguna reflexión previa explícita y sin la necesidad de un autocontrol firme. Otras son inducidas por la maldad cuando quieren ser buenas y por la bondad cuando tienen intenciones de ser malas, también sin ninguna reflexión previa y sin una resistencia activa a estas violaciones de sus deseos de orden superior. Gozar de la libertad les es fácil a algunos. Otros deben luchar por lograrlo.

IV

Mi teoría respecto de la libertad de la voluntad explica fácilmente nuestra reticencia a aceptar que los miembros de cualquier especie inferior a la nuestra puedan disfrutar de esta libertad. Asimismo, cumple otra condición que cualquier teoría de este tipo debe satisfacer, evidenciando por qué debe considerarse deseable

la libertad de la voluntad. Disfrutar de libre albedrío supone la satisfacción de ciertos deseos –deseos del segundo orden o de órdenes superiores–, mientras que su ausencia significa su frustración. Las satisfacciones en juego son aquellas que se le confieren a una persona de quien se puede decir que tiene voluntad propia. Las frustraciones correspondientes son aquellas que sufre una persona de quien se puede decir que está alejada de sí misma, o que se considera un espectador indefenso o pasivo frente a las fuerzas que lo inducen.

Una persona que es libre de hacer lo que quiere hacer puede, a pesar de ello, estar privada de la voluntad que quiere. Supongamos, sin embargo, que goza tanto de libertad de acción como de libertad de la voluntad. Entonces, no sólo es libre de hacer lo que quiere hacer; también es libre de querer lo que quiere querer. Me parece que, en ese caso, tiene toda la libertad que es posible desear o concebir. Hay otras cosas buenas en la vida, y es posible que no posea algunas de ellas. Pero no carece de nada en materia de libertad.

No es evidente en absoluto que algunas otras teorías sobre la libertad de la voluntad satisfagan estas condiciones elementales pero básicas: que sea comprensible por qué deseamos esta libertad y por qué rechazamos atribuírsela a los animales. Consideremos, por ejemplo, la curiosa versión de Roderick Chisholm de la doctrina de que la libertad humana implica una ausencia de determinación causal.[7] Cuando una persona lleva a cabo una acción libre, según Chisholm, se trata de un milagro. El movimiento de la mano de una persona, cuando ésta la mueve, es el resultado de una serie de causas físicas; pero algún acontecimiento de esta serie, "y presumiblemente uno de los que se produjeron en el cerebro, fue causado por el agente y por ningún otro acontecimiento" (p. 18). Un agente libre tiene, por tanto, "una prerrogativa que algunos sólo le atribuirían a Dios: cada uno de nosotros, cuando actuamos, es la principal fuerza motriz impasible" (p. 23).

Esta explicación no ofrece ningún fundamento para dudar de que los animales de especies subhumanas gocen de la libertad que

7 R. Chisholm, "Freedom and Action", en K. Lehrer (ed.), *Freedom and determinism*, Nueva York, Random House, 1966, pp. 11-44.

define. Nada de lo que dice Chisholm nos permite considerar menos probable el carácter milagroso de la acción de un conejo al mover la pata que el de la acción de un hombre al mover la mano. Pero, en todo caso, ¿por qué le debería *importar* a alguien la posibilidad de interrumpir el orden natural de las causas de la manera que describe Chisholm? Chisholm no da razones para creer que existe una diferencia discernible entre la experiencia de un hombre que inicia milagrosamente una serie de causas cuando mueve la mano y un hombre que mueve la mano sin violar la secuencia causal normal. No parecería haber fundamentos concretos para preferir estar envuelto en una situación más que en la otra.[8]

Por lo general se supone que, además de cumplir con las dos condiciones que he mencionado, una teoría satisfactoria de la libertad de la voluntad necesariamente proporciona un análisis de una de las condiciones de la responsabilidad moral. El abordaje reciente más común para el problema de cómo comprender la libertad de la voluntad ha sido, de hecho, indagar qué implica la suposición de que alguien es moralmente responsable de lo que hizo. No obstante, en mi opinión, la relación entre responsabilidad moral y la libertad de la voluntad ha sido, por lo general, mal entendida. No es verdad que una persona sea moralmente responsable de lo que hizo sólo si su voluntad era libre cuando lo hizo. Es posible que sea moralmente responsable de haberlo hecho incluso si su voluntad no era en absoluto libre.

La voluntad de una persona es libre sólo si esa persona es libre para tener la voluntad que quiere. Esto significa que, respecto de cualquiera de sus deseos de primer orden, es libre ya sea de hacer que ese deseo sea su voluntad, ya de hacer que algún otro deseo de primer orden sea su voluntad. Sea cual fuere su voluntad, entonces, la voluntad de la persona cuya voluntad es libre podría haber sido otra; podría haber hecho otra cosa diferente que constituir su voluntad tal como lo hizo. La cuestión de cómo entender "podría

8 No estoy sugiriendo que la supuesta diferencia entre estas dos situaciones sea inverificable. Por el contrario, los fisiólogos bien podrían demostrar que las condiciones impuestas por Chisholm para definir una acción libre no se cumplen, estableciendo que no hay ningún acontecimiento mental pertinente para el cual no se pueda encontrar una causa física suficiente.

haber hecho otra cosa" en contextos como éste resulta polémica. Sin embargo, aunque esta cuestión es importante para la teoría de la libertad, no tiene relación con la teoría de la responsabilidad moral, ya que la suposición de que una persona es moralmente responsable de lo que hizo no implica que la persona estuviera en una posición de tener la voluntad que quería.

Lo que esta suposición sí implica es que la persona hizo lo que hizo libremente, o que lo hizo por su propio libre albedrío. No obstante, es un error creer que alguien actúa libremente sólo cuando es libre de hacer lo que quiere o que actúa según su libre albedrío si su voluntad es libre. Supongamos que una persona hizo lo que quería hacer, que lo hizo porque quería hacerlo, y que la voluntad por la cual fue inducida cuando lo hizo era su voluntad, porque era la voluntad que quería. Entonces, lo hizo libremente y por su propio libre albedrío. Aun en el supuesto de que hubiera podido hacer otra cosa, no habría querido hacer otra cosa; y aun en el supuesto de que hubiera podido tener una voluntad diferente, no habría querido que su voluntad difiriera de la que era. Además, debido a que la voluntad que la indujo cuando actuó era su voluntad porque ella quería que así lo fuera, no puede alegar que fue obligada a tener esa voluntad ni que asistió pasivamente a su constitución. En estas condiciones, es bastante irrelevante para la evaluación de su responsabilidad moral preguntar si realmente tenía a su disposición las alternativas que descartó.[9]

A manera de ejemplo, consideremos un tercer tipo de adicto. Supongamos que su adicción tiene la misma base fisiológica y la misma embestida irresistible que las adicciones de los adictos contra su voluntad y los adictos inconscientes, pero que está totalmente encantado con su condición. Es un adicto por voluntad propia, que no querría que las cosas fueran distintas. Si la fuerza de su adicción en cierta forma se debilitara, haría todo lo que estuviera a su alcance para reavivarla; si su deseo por la droga comenzara a mermar, tomaría medidas para renovar su intensidad.

9 Véase en "Posibilidades alternativas y responsabilidad moral", capítulo 1 de este volumen, otro análisis de las consideraciones que ponen en duda el principio de que una persona es moralmente responsable de lo que hizo sólo si podría haber hecho otra cosa.

La voluntad del adicto por voluntad propia no es libre, puesto que su deseo de consumir la droga será efectivo independientemente de si quiere o no que este deseo constituya su voluntad. Pero cuando consume la droga, lo hace libremente y por propio libre albedrío. Estoy inclinado a interpretar su situación como sobredeterminada por su deseo de primer orden de consumir la droga. Este deseo es su deseo efectivo porque fisiológicamente es adicto. Sin embargo, también es su deseo efectivo porque él quiere que lo sea. Su voluntad está fuera de su control, pero, debido a su deseo de segundo orden de que su deseo de la droga sea efectivo, ha convertido a esta voluntad en propia. Debido a que su deseo de la droga es efectivo no sólo por su adicción, es posible que sea moralmente responsable de consumirla.

Mi idea de la libertad de la voluntad es, en apariencia, neutral respecto del problema del determinismo. Se puede pensar que el hecho de que una persona sea libre de querer lo que quiere querer esté causalmente determinado. Si esto es concebible, entonces el que una persona goce de libre albedrío también podría estar determinado causalmente. No hay más que una inocua apariencia de paradoja en la proposición que sostiene que está establecido, en forma inevitable y por fuerzas más allá de su control, que ciertas personas tienen libre albedrío y otras no. No hay incoherencia en la proposición de que una agencia, diferente de la propia de la persona, sea responsable (incluso *moralmente* responsable) del hecho de que esa persona disfrute o no de la libertad de la voluntad. Es posible que una persona sea moralmente responsable de lo que hace por su propio libre albedrío, y que alguna otra persona también sea moralmente responsable de que la primera lo haya hecho.[10]

10 Existe una diferencia entre ser *completamente* responsable y ser *únicamente* responsable. Supongamos que el adicto por voluntad propia se ha hecho adicto por el trabajo deliberado y calculado de otro. Entonces, podría ser que tanto el adicto como esta otra persona fueran completamente responsables de que el adicto consuma la droga, mientras que ninguno de ellos es únicamente responsable de ello. El hecho de que existe una distinción entre la responsabilidad moral completa y la responsabilidad moral única es evidente en el siguiente ejemplo. Hay una luz que puede encenderse o apagarse activando uno de dos interruptores, y cada uno de estos interruptores es activado simultáneamente a la posición de "encendido" por dos personas diferentes que desconocen la existencia

Por otra parte, podemos imaginar que podría suceder por casualidad que una persona fuera libre de tener la voluntad que quiere. Si cabe la posibilidad de que esto suceda, podría ser cuestión de suerte que algunas personas gocen de la libertad de la voluntad, y que otras no lo hagan. Quizá también se pueda imaginar, como creen varios filósofos, que las situaciones acontecen de manera no casual o como resultado de una secuencia de causas naturales. Si en realidad se puede concebir que las situaciones pertinentes acontezcan de una manera diferente de las dos anteriores, también es posible que una persona, en esa tercera manera, llegue a gozar de la libertad de la voluntad.

---

de la otra. Ninguna de las personas es únicamente responsable de que la luz esté encendida, ni tampoco comparten la responsabilidad en el sentido de que cada una es parcialmente responsable; más bien, cada una de ellas es completamente responsable.

3
Coacción y responsabilidad moral

> En ciertos casos, si no se llega hasta a alabar, por lo menos se perdona a un hombre que hace lo que no debe en circunstancias superiores a las fuerzas ordinarias de la naturaleza humana, y que nadie podría resistir.
> **Aristóteles, *Ética a Nicómaco* III, 1, 1110a**

I

Los tribunales pueden negarse a admitir como prueba una confesión –alegando que se hizo bajo coacción– que la policía obtuvo de un prisionero bajo la amenaza de golpearlo. Sin embargo, es menos probable que los cómplices del prisionero, que quedaron comprometidos por su confesión, acepten que éste realmente confesó bajo coacción. Es posible que sientan, quizá justificadamente, que él hizo una elección censurable y que actuó mal: debería haber aceptado la paliza en lugar de traicionarlos. Es decir, en ocasiones –aunque no siempre–, el uso del término *coacción* significa la exclusión de la responsabilidad moral. Por esa razón, se considera que una persona que actúa bajo coacción no lo hace libremente o por su propia voluntad. Puede afirmarse que no se puede ni reconocer mérito ni culpar a una persona por lo que ha hecho, entonces, demostrando que fue *coaccionada* a hacerlo.

A veces se dice que una persona ha sido coaccionada aun cuando no ha llevado a cabo ninguna acción. Supongamos que un

hombre aplica una fuerte presión en la muñeca de otro y lo fuerza a soltar el cuchillo que tiene en la mano. En este caso, que involucra lo que puede llamarse una *coacción física*, la víctima no es obligada a actuar; lo que sucede es que sus dedos se aflojan por la presión aplicada en su muñeca. En algunas situaciones, puede ser difícil, o aun imposible, saber si se llevó a cabo una acción o no. Tal vez, no quedará claro si el hombre soltó el cuchillo porque sus dedos se aflojaron o porque deseaba evitar que continuara la presión en su muñeca. O supongamos que están torturando severamente a un hombre para obligarlo a revelar una contraseña, y que en cierto momento pronuncia la palabra. Puede no haber forma de descubrir si dijo la palabra ante la amenaza de sufrir más dolor o si –vencida su voluntad por la agonía de lo que ya había sufrido– la palabra salió involuntariamente de su boca.

Propongo considerar los casos de coacción en los que la víctima es obligada a llevar a cabo una acción, al proporcionársele algún motivo para hacerlo, pero que se asemejan a casos de coacción física por el hecho de que no debe considerarse a la víctima moralmente responsable por lo que ha sido obligada a hacer. Podríamos decir que en casos de coacción física, el cuerpo de la víctima es usado como un instrumento cuyos movimientos están sometidos a la voluntad de otra persona. En cambio, en aquellos casos de coacción que me ocupan, la voluntad de la víctima es sometida a la voluntad de otro. En esos casos, ¿cómo afecta la coacción la libertad de su víctima? ¿Qué fundamento proporciona para juzgar que la víctima no es moralmente responsable por hacer lo que ha sido obligada a hacer?

Hay distintas formas en las que una persona puede intentar motivar a otra a llevar a cabo cierta acción. Limitaré mi atención a sólo dos de ellas: proferir una amenaza condicional y hacer un ofrecimiento condicional. En cada una, una persona (P) propone originar cierta situación (C) si la otra persona (Q) lleva a cabo cierta acción (A). La cuestión de si una persona que hace una propuesta de este tipo en realidad está profiriendo una amenaza o haciendo un ofrecimiento depende, en parte, de sus razones, sus intenciones y sus creencias. Las mismas consideraciones deben hacerse también al interpretar la respuesta subsiguiente de la per-

sona que recibe la amenaza o el ofrecimiento. Sin embargo, para simplificar mi análisis, en general pasaré por alto estos factores. Cuando digo que una persona hace una amenaza o un ofrecimiento, se debe suponer que satisface todas las condiciones necesarias para hacerlo respecto de sus razones, sus intenciones y sus creencias, de la misma manera que cuando digo que una persona se somete a una amenaza o le hace frente, o bien acepta o rechaza un ofrecimiento. También se debe suponer que todas las amenazas y los ofrecimientos son creíbles y firmes: todas las personas involucradas tienen razones suficientes para creer que las propuestas en cuestión serán llevadas a cabo si sus condiciones se cumplen.

Las formulaciones condicionales de amenazas y ofrecimientos suelen ser, en realidad, implícita o explícitamente *bi*condicionales. Cuando P propone provocar C si Q hace A, muchas veces también expresa o insinúa que no provocará C si Q no hace A. Esto no es necesariamente así. P puede dejar abierta la suposición de que cumplirá su amenaza u ofrecimiento aun cuando Q realice una acción diferente de la relacionada con la propuesta de P. Sin embargo, cuando un asaltante de caminos le dice a un viajero que debe elegir entre su dinero o su vida, parecería que la frase debe ser interpretada como bicondicional: el asaltante de caminos matará al viajero si éste se rehúsa a entregarle su dinero, mientras que, en caso contrario, le perdonará la vida. Y cuando un empleador ofrece pagarle a alguien cierto salario por hacer cierto trabajo, al potencial empleado suele quedarle claro que no se le pagará el salario si declina el ofrecimiento.[1]

Puede parecer razonable interpretar que cada propuesta bicondicional que expresa una amenaza o un ofrecimiento necesariamente expresa ambos a la vez. Cuando el asaltante de caminos amenaza con matar al viajero, se puede pensar también que está ofreciendo al viajero su vida a cambio de su dinero; y cuando el empleador le ofrece un puesto a alguien, puede parecer que está amenazando implícitamente con retener el dinero de esa persona

1 No analizaré propuestas complejas, en las que P propone provocar cierta consecuencia si Q hace A, y hacer algo más que abstenerse de provocar esa consecuencia si Q no hace A.

a menos que acepte el puesto en cuestión. Sin embargo, estoy de acuerdo con Nozick en rechazar la interpretación de que cualquier propuesta bicondicional que haga una amenaza o un ofrecimiento también está haciendo un ofrecimiento o una amenaza correspondientes.[2] Sin duda, por lo general, no se considera que un comerciante amenaza a sus clientes –ni siquiera mediante insinuaciones– cuando les ofrece sus productos en venta, aunque su ofrecimiento de vender esté naturalmente combinado con una propuesta de retener los productos si el cliente no acepta pagar su precio.

Las amenazas y los ofrecimientos difieren de diversos modos. Una persona que cumple con las condiciones de un ofrecimiento suele tener la opción de no aceptar lo que se le ha ofrecido a cambio, mientras que, habitualmente, esta opción no existe para alguien que cumple con las condiciones de una amenaza. Puede resultar sensato que una persona que ha recibido un ofrecimiento siga mirando para comprar algo mejor, pero alguien que ha recibido una amenaza no tiene una alternativa sensata equivalente. En general, se piensa que amenazar a una persona requiere una justificación, mientras que no existe una presunción similar contra la legitimidad de hacerle un ofrecimiento a alguien.

La diferencia fundamental entre las amenazas y los ofrecimientos, no obstante, es la siguiente: una amenaza pone al destinatario en peligro de recibir un castigo, mientras que un ofrecimiento le brinda la posibilidad de obtener un beneficio. Si la mitad de una propuesta bicondicional es una amenaza, entonces, la otra mitad sería un ofrecimiento si y sólo si el hecho de no infligir el castigo con el cual se amenazó equivaliera a conceder un beneficio; y si una de las mitades de una propuesta bicondicional es un ofrecimiento, la propuesta que la acompaña sería una amenaza si y sólo si el hecho de no brindar el beneficio ofrecido equivaliera a impo-

2 R. Nozick, "Coercion", en S. Morgenbesser, P. Suppes y M. White (comps.), *Philosophy, science, and method: essays in honor of Ernest Nagel*, Nueva York, St. Martin's Press, 1969, p. 447. Realmente estoy en deuda con este espléndido ensayo, que ha proporcionado una base indispensable para mi propio análisis de algunos de los temas que trata. Aunque tengo una opinión crítica con respecto a algunas de las opiniones de Nozick, mi ensayo sigue al suyo en distintos aspectos, que resultarán evidentes a cualquiera que esté familiarizado con ambos.

ner un castigo. Pero, ¿cuáles son las características de los castigos y los beneficios, y en qué condiciones abstenerse de uno equivale a imponer o conceder el otro? ¿Qué determina si la propuesta de P de provocar C si y sólo si Q hace A incluye un ofrecimiento, o sólo una amenaza, o ambos a la vez?

Nozick sugiere el siguiente criterio para distinguir las amenazas de los ofrecimientos: si C "hace que las consecuencias de la acción de Q sean peores de lo que habrían sido en el curso normal y esperado de los acontecimientos", entonces la propuesta de P es una amenaza; si C hace que las consecuencias sean mejores, la propuesta es un ofrecimiento. Explica a continuación que "se pretende que el término *esperado* se mueva o esté a horcajadas entre *previsto* y *moralmente requerido*".[3] Ahora bien, este criterio requiere que el curso de los acontecimientos cuando Q hace A y P provoca C pueda compararse con otro curso de los acontecimientos en que Q hace A. Sin embargo, no está del todo claro cómo se puede identificar este segundo curso de los acontecimientos, que proporciona la base para evaluar la importancia de la propuesta de P. ¿Cuáles son las consecuencias *normales y esperadas* cuando Q hace A, con las que se deben comparar las consecuencias que se producen cuando P provoca C?

El criterio de Nozick admite una variedad de interpretaciones. Consideremos la interpretación que él mismo le da cuando la aplica en su análisis de las siguientes dos situaciones:

(1) P es el proveedor habitual de drogas de Q y hoy, cuando va a verlo a Q, dice que no se las venderá, como lo hace normalmente, por veinte dólares, sino que, en cambio, se las entregará si y sólo si Q le da una paliza a cierta persona.

(2) P es un extraño que ha estado observando a Q y sabe que Q es drogadicto. Ambos saben que el proveedor habitual de drogas de Q fue arrestado esta mañana y que P no tuvo nada que ver con su arresto. P se acerca a Q y le dice que le proveerá drogas si y sólo si Q le da una paliza a cierta persona.[4]

3 *Ibid.*, p. 447. Nozick no hace comentario alguno acerca de la distinción entre *normal* y *previsto*.

4 *Ibid.*

Nozick cree que la segunda de estas situaciones no contiene amenaza alguna, sino sólo un ofrecimiento. "En el curso normal de los acontecimientos", explica, P_2 "no provee drogas a Q en absoluto, y tampoco se espera que lo haga." Si P_2 no le da drogas a Q, "no le está *reteniendo* drogas a Q ni tampoco está *privando* a Q de las drogas". Simplemente le está "ofreciendo drogas a Q", por tanto, "como un incentivo para que golpee a otra persona". En cambio, Nozick sostiene que la primera situación contiene tanto un ofrecimiento como una amenaza. Dado que, "en el curso normal de los acontecimientos, [P_1] le provee drogas a Q a cambio de dinero", los términos de esta propuesta significan que P_1 empeorará la situación de Q en caso de que Q no golpee a la persona. Por tanto, P_1 está amenazando a Q. Por supuesto, también le está haciendo un ofrecimiento: dado que Q normalmente no recibe drogas de P_1 por golpear a una persona, la conveniencia para Q de realizar esta acción está intensificada por la propuesta de P_1.[5]

En mi opinión, lo que Nozick dice acerca de estas dos situaciones es erróneo. La propuesta de P_1, como él sostiene, sin duda incluye tanto una amenaza como un ofrecimiento. Pero el criterio que emplea para identificar las amenazas y los ofrecimientos lo lleva a dar una explicación incorrecta de por qué la propuesta tiene este carácter dual. Más aun, cuando este criterio se reemplaza por uno más satisfactorio, resulta evidente que Nozick también se equivoca al considerar la propuesta de P_2 sólo como un ofrecimiento. El hecho es que la propuesta de P_2 incluye no sólo un ofrecimiento, sino también una amenaza. Y lo hace en virtud de las mismas características que hacen que parte de la propuesta de P_1 sea una amenaza.

Si la propuesta de P_1 debiera interpretarse –por la razón que menciona Nozick– como una amenaza de castigo para Q, entonces un carnicero estaría amenazando a sus clientes con un castigo cada vez que subiera el precio de la carne. Lo que P_1 hace al sustituir su antigua propuesta por la nueva es, después de todo, simplemente subir el precio de la droga. En lugar de pedirle a Q veinte dólares por ella, como antes, ahora le pide que haga algo para obtener la

5 R. Nozick, "Coercion", *op. cit.*, pp. 447-448.

droga que (debemos suponer) a Q le gusta menos que darle veinte dólares a P_1. Ahora bien, sin duda el carnicero no está proponiendo castigar a sus clientes sólo porque les dice que va a cambiar el precio de un modo que resultará desventajoso para ellos. Es probable que, al decir esto, continúe haciendo sólo un ofrecimiento, aunque menos atractivo que el anterior.[6] Así, el hecho de que P_1 empeore la situación del drogadicto al cambiar los términos de su propuesta de venderle drogas no puede ser, como afirma Nozick, lo que explica el hecho de que la propuesta de P_1 sea una amenaza. El criterio de Nozick, como él mismo lo interpreta, es inaceptable: no preserva la distinción entre amenazar con castigar a alguien y dificultar sus opciones haciéndole un ofrecimiento peor que el anterior.

Para decidir si P_1 y el carnicero están haciendo amenazas u ofrecimientos, no es esencial comparar, como recomienda Nozick, lo que proponen hacer con lo que *solían* hacer. Más bien, es necesario comparar los acontecimientos que anticipan sus propuestas actuales con lo que ocurriría *ahora* si ellos no intervinieran del modo en que se proponen hacerlo. Y la pregunta sobre qué sucedería ahora sin estas intervenciones no debe responderse citando los términos de las propuestas anteriores del carnicero y de P_1 puesto que éstos han quedado enteramente anulados por los términos que los han reemplazado. Cuando hacen sus nuevas propuestas, P_1 y el carnicero, en realidad, hacen dos cosas: borran sus propuestas anteriores y establecen nuevos términos. Dado que los términos de las antiguas propuestas se borran y se comienza de cero, no hay un fundamento para suponer que las propuestas anteriores sirven

6 Puede objetarse que, a pesar de mi deseo de dejar a un lado estas consideraciones, las razones del carnicero deben tomarse en cuenta. Sin embargo, aunque supongamos que parte del objetivo del carnicero al aumentar sus precios sea empeorar la situación de sus clientes (véase la condición 3' de Nozick, *op. cit.*, p. 442), no se desprende de ello que su propuesta sea una amenaza. La propuesta puede continuar siendo simplemente un ofrecimiento, a pesar de ser (intencionalmente) tan poco atractiva como para disuadir a los clientes del carnicero de seguir comprándole. Por supuesto, la propuesta del carnicero *puede* ser una amenaza. Pero si lo es, no es porque (intencionalmente) sea menos favorable para los clientes que una propuesta anterior. Lo es por otras circunstancias, que analizaré más adelante.

para definir lo que sucedería ahora si los términos de las propuestas actuales no se llevaran a cabo, que es lo único que justificaría el procedimiento de Nozick.

No estoy sugiriendo que los términos en los que la gente ha hecho tratos en el pasado no tengan relación alguna con la evaluación de los términos que gobiernan sus tratos actuales. Aquellos términos anteriores, como explicaré más adelante, pueden tener relevancia en el presente. Más aun, no estoy sugiriendo que los términos de una propuesta anterior nunca definan el punto de partida apropiado para medir los acontecimientos previstos por una propuesta actual. Imaginemos que un fiscal dice que pedirá la pena capital si el acusado se declara inocente, y que más tarde propone pedir una pena menor aunque el acusado se declare inocente *si* el acusado le ofrece suficientes pruebas contra otra persona. Aquí, la propuesta anterior del fiscal define el punto de partida para evaluar su segunda propuesta; sigue siendo decididamente pertinente para la pregunta de qué sucedería si el acusado rechazara la última propuesta. Pero esto es porque, a diferencia de lo que sucede en las situaciones que involucran al adicto y al cliente del carnicero, la segunda propuesta del fiscal no anula por completo los términos de la primera.

A fin de evaluar una propuesta de P de intervenir en la serie de acontecimientos que se inicia cuando Q hace A, debemos saber si esta intervención de P dejará a Q en una situación mejor o peor de la que tendría sin esa intervención. Medir el impacto de la propuesta requiere, por tanto, que se compare el curso de los acontecimientos cuando P interviene de acuerdo con los términos de su propuesta con lo que sucederá si esta intervención se quita de esos acontecimientos. Esta comparación excluye los términos de cualquier propuesta que sea anulada por la propuesta actual de P y toma en cuenta los términos de cualquier propuesta que permanezca intacta.[7]

7 Supongamos que P ofrece pagarle a Q cierta cantidad de dinero por ir a trabajar para él. El criterio de Nozick nos haría comparar esto con las consecuencias *normales y esperadas* de la decisión de Q de trabajar para P. Sin embargo, ¿qué es *normal* o *esperable* cuando Q va a trabajar para P? Quizá la mejor respuesta es que P le pague a Q un salario justo por su trabajo. No obstante, claramente la

Proceder de esta manera evita el error de Nozick y es correcto hasta cierto punto. Pero no nos lleva al final del asunto, puesto que no nos permite tratar en forma satisfactoria las situaciones en las que P propone *no* intervenir de cierta manera en la serie de acontecimientos iniciada por la acción de Q. Darle una paliza a una persona y obtener drogas de P_1 es más deseable, desde el punto de vista del adicto, que hacer lo mismo sin obtener drogas de P_1. Es decir que la propuesta de P_1 de darle drogas al adicto si y sólo si éste le da una paliza a alguien incluye un ofrecimiento, aunque no sea un ofrecimiento tan bueno como aquel a partir del cual P_1 y el adicto hacían negocios antes. La propuesta del carnicero a su cliente incluye un ofrecimiento análogo; es mejor para el cliente obtener carne a cambio de su dinero –aunque sea menos carne que la que obtenía antes por la misma cantidad de dinero– que darle al carnicero el dinero y no recibir carne. Pero supongamos que el adicto se niega a darle una paliza a alguien y que el cliente se niega a pagarle al carnicero.

Entonces, el curso de los acontecimientos previsto por las dos propuestas es: P_1 no le da drogas al adicto y el carnicero no le da carne al cliente. En los casos en cuestión, P_1 y el carnicero proponen no intervenir en absoluto. No agregarán ni quitarán nada a la serie de acontecimientos que ocurrirían si ellos no fueran conscientes de lo que el adicto y el cliente hacían o no hacían, y si no respondieran en absoluto a sus acciones. No obstante, sería un error llegar a la conclusión de que sus propuestas –no darle drogas al adicto ni carne al cliente del carnicero– son de la misma naturaleza. Por el contrario, P_1 amenaza al adicto cuando afirma que no le dará drogas si éste se niega a darle una paliza a la persona, mientras que el carnicero no expresa una amenaza semejante. ¿Cómo debemos explicar el hecho de que una propuesta de abstenerse de brindar cierto beneficio es una amenaza en un caso y no en el otro?

---

comparación apropiada no debe plantearse entre lo que P dice que va a pagarle a Q y lo que sería justo pagarle a Q (o lo que la gente normalmente cobra por ese trabajo, o lo que Q cobra normalmente, o lo que Q normalmente cobra de P, o lo que P normalmente paga, o lo que P normalmente le paga a Q). Debe plantearse entre el hecho de que P le dé a Q lo que dice que le dará y el que P no le dé a Q nada a cambio de su trabajo.

Sería útil enfocar este problema concentrando la atención en ciertas cosas que tendemos a dar por sentadas al pensar en las situaciones en las que participan P_1 y el carnicero. Consideremos entonces cómo interpretaremos la propuesta bicondicional que P_1 le hace al adicto si suponemos que existe una enorme sobreoferta de drogas en el mercado y que el adicto tiene un acceso cómodo a numerosos vendedores, cuyos precios son bastante más bajos que los de P_1. Consideremos también cómo interpretaremos la propuesta del carnicero si suponemos que el cliente se morirá de hambre si el carnicero no le da carne y que el precio del carnicero es escandalosamente alto. Cambiar de esta manera nuestros supuestos con respecto a las dos situaciones nos llevaría, creo, a modificar nuestras evaluaciones. Ya no consideraríamos que P_1 está amenazando al adicto, sino sólo haciéndole un ofrecimiento poco atractivo. Al mismo tiempo, interpretaríamos que la propuesta del carnicero a su cliente incluye no sólo un ofrecimiento sino también una importante amenaza.

No hallamos elementos de amenaza en la propuesta del carnicero de subir sus precios siempre que supongamos que, al hacer esta propuesta, no se está aprovechando en forma indebida de la situación en la que tiene al cliente en sus manos. Su propuesta adquiere el carácter de una amenaza, por otra parte, cuando se cumplen tres condiciones. Primero, el cliente *depende* del carnicero para obtener carne: no puede conseguirla fácilmente de otra fuente. Segundo, el cliente *necesita* carne: ella es esencial para evitar lo que él consideraría un importante deterioro de su bienestar o para dejar de estar en una situación que él considera no deseable. Tercero, el carnicero *explota* la dependencia y la necesidad del cliente; por su carne exige un precio injusto o inadecuado. Cuando se cumplen las dos primeras condiciones, el carnicero tiene al cliente en su poder. Si luego ofrece carne a un precio abusivo, su propuesta de abstenerse de darle carne al cliente si éste no le paga lo que le pide constituye una amenaza.

Es poco razonable considerar que la propuesta de P de negarle a Q cierto beneficio equivale a una amenaza –aun a una amenaza débil o poco efectiva– de castigarlo, a menos que Q no pueda obtener con facilidad un beneficio equivalente en otra parte. Pues

sólo en ese caso Q tiene una razón para estar interesado en la cuestión de si obtiene el beneficio de P o no, y un castigo al que es razonable ser enteramente indiferente no es un castigo en absoluto.[8] Con respecto a la segunda de las condiciones que he especificado, supongamos que P propone darle a Q un millón de dólares si y sólo si Q lleva a cabo cierta acción, que Q no tiene otra oportunidad de adquirir tanto dinero, y que el ofrecimiento de P es, de alguna manera, injusto o indebido. Aun así, la propuesta no incluye una amenaza porque (supongamos) la continuidad del bienestar de Q por encima de un nivel que él considera poco deseable no depende de que él tenga un millón de dólares. Si bien es posible que desee el dinero con desesperación una vez que la propuesta de P le haga sentir que está a su alcance, no lo necesita. La cuestión central de la tercera condición es que, sin ella, alguien que tuviera el control monopólico de una necesidad estaría profiriendo una amenaza si pidiera cualquier precio por el beneficio que tiene bajo control. Sin embargo, sería poco razonable interpretar que los proveedores de electricidad estarían amenazando al público, a pesar de que poseen el control monopólico de una necesidad, incluso en el caso de que se propusieran vender la electricidad a un precio filantrópico, un precio muy por debajo de su propio costo, digamos, y dentro de las posibilidades de todos.

Al considerar si el precio establecido para determinado beneficio resulta abusivo, puede ser apropiado tener en cuenta el precio al que beneficios similares han sido otorgados en el pasado. Además, la historia de la relación entre dos personas puede arrojar luz sobre la cuestión de si una de ellas depende de la otra para alguna necesidad. Más aun, el hecho de que una persona habitualmente haya deseado efectuar transacciones con otra a un precio dado puede, bajo ciertas condiciones, crearle la obligación de continuar

8 Supongamos que el cliente pudiera, en efecto, obtener carne comparable en otra carnicería cercana, pero sólo al mismo precio ridículamente alto o, incluso, más caro. En ese caso, el carnicero también estaría expresando una amenaza al subir el precio –lo hubiera hecho en connivencia con los demás o no–, porque el cliente no tendría, de hecho, ninguna opción útil fuera de las que define la propuesta del carnicero. Su dependencia del carnicero, por tanto, no está mitigada en forma significativa por el hecho de tener otras alternativas. Pero quizá debería decirse que, en virtud de esas otras alternativas, el carnicero no lo tiene tan en *su* poder.

manteniendo el mismo precio aunque fuera justo y correcto que otro estableciera un precio más alto. Quizá también haya otras formas en las que el pasado pueda tener un papel significativo en relación con la cuestión de si la contrapartida de un ofrecimiento actual es una amenaza. Pero ninguna de estas consideraciones justifica la afirmación de Nozick de que, mientras que la propuesta que P_1 le hace al adicto es tanto un ofrecimiento como una amenaza, la propuesta de P_2 sólo es un ofrecimiento.

La propuesta de P_1 no incluye una amenaza a menos que supongamos que el adicto depende de él para obtener drogas, que el adicto necesita drogas y que el precio que P_1 pide es abusivo. Parece tan razonable hacer estas suposiciones con respecto a la segunda de las dos situaciones de Nozick como hacerlas con respecto a la primera. La propuesta de P_2 es lisa y llanamente un ofrecimiento, ya que prevé una intervención beneficiosa de P_2 en los acontecimientos, que serían menos atractivos para el adicto sin esta intervención. Pero la propuesta también incluye una amenaza, por la misma razón por la que hay una amenaza en la propuesta de P_1.

Con su proveedor habitual en manos de la policía, el adicto reconocerá que depende de P_2 para obtener drogas en cuanto P_2 revele su disposición para proveérselas. El hecho de que P_2 no le haya dado drogas al adicto en el pasado no es pertinente para la cuestión de si P_2 está expresando una amenaza o un ofrecimiento.[9] Lo que cuenta es que P_2 tiene al adicto en su poder al hacerle entender que debe escoger entre hacer lo que P_2 le pide y quedarse sin las drogas que necesita. Además, no hay razón para pensar que el precio en cuestión es más justo o adecuado cuando P_2 les pone precio a las drogas que cuando lo hace P_1. Lo que hace que este precio sea injusto o indebido cuando P_1 lo pide no es el hecho de que sea más alto de lo que suele pedir por las drogas –los aumentos de precios no son necesariamente sinónimo de abuso–, sino que exige que el adicto lleve a cabo una acción injusta y peligrosa. La acción tiene estas mismas características, por supuesto, cuando es exigida por P_2. Suponiendo que P_1 se aprovecha del adicto al aumentar el

9 Este hecho, no obstante, puede significar que P_2 actúa de un modo menos censurable que P_1.

precio, entonces, P_2 también se aprovecha de él cuando establece ese precio más alto por su transacción inicial.

El hecho de no brindar un beneficio es, en las condiciones que acabo de bosquejar, equivalente a imponer un castigo. ¿En qué condiciones el no infligir un castigo otorga un beneficio? Supongamos que alguien ha robado cinco mil dólares, que ha gastado todo menos cien dólares, que ha sido declarado culpable por su delito, y que el juez propone enviarlo a la cárcel por diez años si y sólo si se niega a pagar una multa de cincuenta dólares. Aquí el juez está amenazando con castigar al delincuente enviándolo a la cárcel si no paga la multa, pero también le está ofreciendo al criminal un beneficio –su libertad– si la paga. La propuesta del juez ofrece un beneficio porque el precio que establece por no cumplir su amenaza es muy bueno. La propuesta le da al delincuente la oportunidad de obtener una ganga –que sería lo opuesto del abuso (evidentemente no existe un antónimo de *abusado*)– ya que el precio que pide por su libertad está por debajo de lo que habría sido justo y correcto pedir por ella. Generalizando, la propuesta de P de no infligir a Q un castigo con el que lo había amenazado equivale al ofrecimiento de un beneficio si P tiene a Q en su poder en lo que al castigo se refiere –por ejemplo, Q no cuenta con medios inmediatos para evitar el castigo excepto en los términos de P, y el castigo lo privaría de algo que él necesita– y si el precio de P por no infligir el castigo es menor al precio que sería justo y correcto que P exigiera. Así como P abusa de Q cuando se aprovecha injustamente del hecho de que tiene a Q en su poder, también P beneficia a Q cuando se aprovecha menos de su poder sobre él de lo que sería justo y correcto.

Existen tres formas en las que P puede castigar a Q por hacer A. Primero, P puede intervenir en la serie de acontecimientos que se inicia cuando Q hace A y agrega a la serie algo que no habría existido de no ser por su intervención y que hace que la serie de acontecimientos resultante sea menos deseable para Q de lo que habría sido sin esta intervención de P.

Segundo, puede intervenir en la serie de acontecimientos que se inicia cuando Q hace A suprimiendo algo que la serie habría incluido de no haber sido por su intervención y cuya ausencia hace

que la serie de acontecimientos resultante sea menos deseable para Q de lo que habría sido sin la intervención de P. Esto implica, aunque no nos guste, que P amenaza con castigar a Q por hablar cuando dice que apagará su audífono si Q dice otra palabra. El ejemplo es de Nozick, y podríamos tratar de salir adelante adaptando una sugerencia suya: la intervención de P impone un castigo a Q sólo si deja a Q en una situación peor, después de haber hecho A, de lo que habría estado Q si no hubiera hecho A y P no hubiera intervenido.[10] Sin embargo, invocar este criterio significaría que cada vez que P amenaza con castigar a Q por hacer A, es necesariamente mejor para Q abstenerse de hacer A que hacerlo y sufrir el castigo. Parece poco deseable incorporar esto a la noción de castigo; es evidente que algunos castigos son poco efectivos. Estoy dispuesto a aceptar la inferencia de que, en el ejemplo del audífono, P castiga a Q al apagar su audífono. Después de todo, "Dejé de hablar porque me amenazó con apagar el audífono" parece, por lo menos, ligeramente aceptable.

Tercero, P puede –sin agregar ni quitar nada de la serie de acontecimientos que se inicia con la acción A de Q– hacer, de manera injusta o incorrecta, que como consecuencia de la acción de Q, Q no obtenga algo que necesita. Resultaría intolerablemente artificial decir que el adicto recibe drogas de P_2 cuando se niega a darle una paliza a alguien excepto en el caso en que P_2 interviene suprimiendo de la serie de los acontecimientos posteriores la posibilidad de que el adicto las obtenga. El castigo que P_2 inflige al adicto por negarse a golpear a alguien no se realiza mediante una *intervención* de esta clase. Más bien, el adicto no obtiene las drogas porque P_2 convierte este hecho en una consecuencia de que el adicto se haya negado a golpear al otro. Esto es equivalente a adjudicar a las drogas el precio abusivo que significa que el adicto golpee a alguien. Tener que prescindir de las drogas no sería una *consecuencia* del rechazo del adicto a golpear al otro (aunque podría ser una *secuela* de ello) si P_2 no la convirtiera en tal. De acuerdo con esto, es posible interpretar el hecho de que P_2 se abstenga de beneficiar o castigar a Q como un tipo particular de intervención suya –poco

10 R. Nozick, *op. cit.*, p. 443.

visible, por así decirlo– en la serie de acontecimientos que se inicia cuando Q hace A.

También hay tres formas, que se corresponden con éstas de un modo que debería resultar claro, en que P puede otorgar un beneficio a Q por hacer A.

Ofrecerle a alguien un beneficio por llevar a cabo cierta acción aumenta para esa persona la conveniencia de hacerla, mientras que amenazarlo con castigarlo por llevarla a cabo reduce la conveniencia que ésta representa para él. Un ofrecimiento (amenaza) será *superfluo* (*/a*) si aumenta (reduce) la conveniencia de una acción que ya es más (menos) deseable que su alternativa. Será *infructuosa* si aumenta (reduce) la conveniencia de una acción sin lograr hacerla más (menos) deseable que su alternativa. Las amenazas y los ofrecimientos superfluos e infructuosos son, por supuesto, amenazas u ofrecimientos a pesar de todo.

II

Pero ¿qué es la coacción? Ejercer coacción sobre alguien para que lleve a cabo cierta acción no puede ser, si su consecuencia ha de ser liberarlo de responsabilidad moral, simplemente una cuestión de lograr que lleve a cabo la acción mediante una amenaza. Una persona coaccionada es *obligada* a hacer lo que hace. No tiene más *opción* que hacerlo. Esto es, por lo menos, parte de lo que es esencial si la coacción ha de relevar a su víctima de la responsabilidad moral, si ha de volver inapropiado tanto elogiarla como culparla por haber hecho lo que fue coaccionada a hacer. Ahora bien, no es necesariamente verdad que una persona que decide evitar un castigo con el que fue amenazada esté obligada a hacerlo o que no tenga otra opción. Tampoco es verdad que una persona no sea moralmente responsable de lo que ha hecho sólo porque lo hizo sometida a una amenaza. Puede decirse que esta persona está actuando "bajo presión"; aunque no toda presión es una coacción.

Podría sugerirse que alguien es coaccionado si, además de su conducta para evitar un castigo producto de una amenaza, se

cumplen otras dos condiciones: (1) el castigo con el que es amenazado hace que la acción contra la que se expresa la amenaza sea *sustancialmente* menos atractiva para él de lo que habría sido en otro caso; y (2) él cree que su situación sería peor si desafiara la amenaza que si se sometiera a ella.[11] Sin embargo, agregar estas condiciones no sirve adecuadamente para identificar los casos de coacción.

Supongamos que P amenaza a Q con pisarlo a menos que Q le prenda fuego a un hospital lleno de gente, y que Q provoca el incendio para que P no lo pise. Esto no satisface la primera condición, que excluye las amenazas triviales: el castigo que Q busca evitar al someterse a las exigencias de P no es sustancial. Supongamos, en cambio, que P amenaza con romperle el pulgar a Q a menos que Q prenda fuego al hospital, y que Q se somete a esta amenaza. Aquí el castigo con el que P amenaza a Q es sustancial: cualquier acción se vuelve sustancialmente menos atractiva para una persona si la consecuencia es un dedo roto de lo que lo sería sin la posibilidad de esta consecuencia. Así, ahora se cumple la primera condición. Más aun, puede ser que la segunda condición también se cumpla. Supongamos que Q piensa que no lo atraparán ni lo castigarán por prenderle fuego al hospital y que no espera tener grandes remordimientos de conciencia por hacerlo. Entonces, bien puede creer que su condición sería peor si hiciera frente a la amenaza de P y terminara con un dedo roto que si hace lo que P le exige. Sin embargo, aunque de esta manera se cumplan ambas condiciones, no parece apropiado decir que Q fue coaccionado para provocar un incendio.

¿Por qué nos resistimos a considerar que Q fue coaccionado aunque supongamos que él cree que sufrirá en forma más sustancial haciendo frente a la amenaza de P que sometiéndose a ella? Una sugerencia sería que se debe a que pensamos que, como Q

11 Estas dos condiciones se basan en las condiciones 2' y 7 de Nozick (*op. cit.*, pp. 442-443), que son los únicos puntos de su lista de las condiciones necesarias y suficientes para la coacción pertinentes para la distinción entre amenazas coercitivas y amenazas que son efectivas, pero no coercitivas. Sin embargo, debe notarse que Nozick no pretende definir las condiciones de la exclusión de la responsabilidad moral. Su uso del término *coacción* difiere, por tanto, del mío.

debe darse cuenta de que es mejor sufrir aunque sea un dedo roto que prenderle fuego a un hospital, no puede creer que someterse a la amenaza de P sea una acción justificada o razonable. En consecuencia, podríamos considerar la posibilidad de revisar la segunda condición y hacer que ésta exija que Q crea que sería poco razonable de su parte desafiar la amenaza de P, o crea que está justificado cuando se somete a ella.

Ahora bien, de cumplirse esta condición revisada, se daría el caso de que una persona que ha sido coaccionada a llevar a cabo cierta acción *cree* que no puede ser *culpada* con razón por haberla realizado. Pero el criterio para la coacción que estamos buscando debe lograr más que esto. Debe asegurar que a una persona coaccionada no se la puede declarar *moralmente responsable en absoluto* de lo que ha sido obligada a hacer. Y esto no se cumpliría ni siquiera fortaleciendo la condición (2) de modo que exigiera que Q creyera *con razón* que está justificado al someterse a la amenaza de P, o que sería poco razonable que le hiciera frente. En efecto, en cualquier versión razonable, para que haya coacción no es necesario que se cumpla la segunda condición; tampoco es suficiente, aun si, además, se cumple la primera condición.

Supongamos que P amenaza con quitarle a Q algo que para Q vale cien dólares a menos que Q le dé algo que Q valúa en cincuenta dólares. El castigo de perder algo que vale cien dólares es sustancial. Más aun, razonablemente podemos suponer que Q cree que su situación será peor si hace frente a la amenaza de P que si se somete a ella, y tiene razón en pensar que estaría justificado someterse a ella. Sin embargo, mientras que Q bien puede elegir someterse a esta amenaza, nada lo *obliga* a hacerlo. La elección entre las alternativas a las que lo enfrenta la amenaza de P depende enteramente de él. Por supuesto, debe escoger entre ellas; debe decidir si hace lo que exige P y escapa del castigo, o si se niega a hacerlo y sufre el castigo. Es libre, sin embargo, para tomar cualquiera de las dos decisiones. Y mientras que puede decidir que es mejor y completamente razonable para él hacer lo que P le exige, podría haber decidido hacer otra cosa. La elección es suya, y no existe un fundamento para afirmar que no tiene plena responsabilidad moral por la decisión que tome.

A veces decimos que una persona no tenía opción cuando la alternativa que eligió era claramente superior a sus otras alternativas. Lo que queremos decir, entonces, es que no tuvo una opción *razonable*, que ninguna otra opción más que la que eligió habría sido razonable. No obstante, no tener opción en este sentido no quiere decir que una persona no merezca crédito o culpa por lo que hace. Sin duda, una persona puede ser digna de elogio por haber hecho una elección claramente razonable. Ahora bien, la coacción requiere algo más especial que esto. Requiere que la víctima de una amenaza no tenga más alternativa que someterse, y por esto debe entenderse no sólo que la persona actuaría razonablemente al someterse –y, por tanto, no debe culpársela por hacerlo–, sino más bien que no es moralmente responsable de su sometimiento.

Este requisito sólo puede cumplirse cuando la amenaza apela a deseos o motivos que están más allá de la capacidad de control de la víctima, o cuando la víctima está convencida de que éste es el caso.[12] Si el deseo o la razón de la víctima para evitar el castigo con el que ha sido amenazada es –o ella considera que es– tan poderoso que no puede impedir que la lleve a someterse a la amenaza, entonces realmente no tiene más alternativa que someterse. Efectivamente, no puede elegir hacer otra cosa. Sólo entonces *puede* ser correcto considerar que carece de responsabilidad moral por su sumisión. El hecho de si, en efecto, *es* adecuado o no considerarla de esa manera –es decir, que fue realmente coaccionada– depende de que se cumpla aun otra condición, que analizaré más adelante.

Una persona puede ser incapaz de hacer frente a una amenaza a la que sabe que sería más razonable hacer frente. Supongamos que alguien tiene un terror patológico de ser picado por una abeja. Puede ser coaccionado a llevar a cabo una acción mediante la amenaza de que, si no lo hace, lo picará una abeja, aunque él mismo reconozca que sería más razonable sufrir la picadura y no llevar a cabo la acción. No estoy afirmando que la coacción se produce sólo

12 No consideraré si la convicción de la víctima debe ser justificada ni de qué manera lo sería. En adelante, para mayor conveniencia, me referiré simplemente a la capacidad de la víctima de controlar sus deseos o sus motivos más que a la condición disyuntiva completa que he formulado antes, suponiendo que la importancia del disyunto faltante resultará obvia.

cuando la víctima de una amenaza es impulsada en forma precipitada a someterse a causa de una ola de pánico. Puede ser coaccionada a hacer lo que se le exige cuando evalúa, con calma, que es incapaz de aceptar el castigo al que se enfrenta.

Tampoco estoy sugiriendo que una persona sea susceptible de ser coaccionada sólo porque tiene, por así decirlo, un repertorio de temores tan imperiosos que puede llegar a hacer cualquier cosa debido a una amenaza que suscita alguno de ellos. La medida en que una persona está en control de sí misma varía en forma considerable de una situación a otra. Un hombre que es fácilmente coaccionado por una amenaza de muerte a entregar su dinero a un ladrón, por ejemplo, puede, sin dudarlo, desafiar la misma amenaza de muerte cuando no es el dinero, sino la vida de su hijo, la que está en juego. Del hecho de que este hombre haya sido capaz de resistirse en la segunda situación no se desprende que no fuera coaccionado en la primera.

No debe suponerse que la diferencia entre estas dos situaciones es simplemente que el hombre juzgó razonable evitar el castigo de muerte en una y que juzgó razonable aceptar ese castigo en la otra. Bien puede haber emitido estos juicios. Pero lo que es esencial hasta el momento, en lo que se refiere a la cuestión de la coacción, es la diferencia en la medida en que pudo, en las dos situaciones, movilizar su fuerza potencial. Advertir que el costo de continuar con determinado curso de acción es mayor que el beneficio hará que una persona piense que continuar con ese curso de acción sería poco razonable. Pero también puede tender a bloquear su acceso a todas sus energías y hacer que le resulte imposible seguir adelante. Saber que lo que se perdería sería demasiado valioso, por otra parte, puede permitirle encontrar recursos en su interior que es incapaz de utilizar en contextos menos graves. Las evaluaciones de una persona pueden afectar no sólo sus juicios con respecto a lo que es razonable hacer. Es posible que también tengan un efecto sobre lo que es capaz de hacer.

Frente a una amenaza coercitiva, la víctima no tiene más opción que someterse: no puede evitar que su deseo de salvarse del castigo en cuestión determine su respuesta. Cuando decide que es razonable someterse a una amenaza no coercitiva, su sumisión no se

vuelve ineludible debido a fuerzas en su interior que es incapaz de doblegar. A veces hablamos de amenazas coercitivas aun cuando no tengamos pruebas específicas de que sus víctimas son incapaces de hacerles frente. Esto se debe a que hay ciertos castigos que no esperamos que alguien elija sufrir. De este modo, una persona que supera esta expectativa no sólo se comporta bien o mal, sino que lo hace con cierta calidad heroica.

En ocasiones, nos parece apropiado emitir un juicio adverso con respecto a la sumisión de una persona a una amenaza, aunque reconozcamos que fue realmente coaccionada y que, por lo tanto, no se la debe considerar moralmente responsable de su sumisión. Esto se debe a que creemos que la persona, aunque de hecho fue incapaz de controlar cierto deseo, debería haber podido hacerlo. Una opinión de esta clase puede estar basada en dos consideraciones, que podrían explicarla. Podemos creer que la persona es moralmente responsable de su propia incapacidad de hacer frente a la amenaza; puede parecernos que es incapaz de resistirse a la amenaza por algo que ella misma ha hecho, y de lo que es moralmente responsable. La otra consideración es sólo una cuestión de juicio moral en un sentido bastante especial. Fundamentalmente es una cuestión de nuestra falta de respeto por la persona que fue coaccionada. Tal vez sea que tenemos una mala opinión de alguien que es incapaz de hacer frente a una amenaza de ese tipo; y nuestra opinión de que debería haber sido capaz de hacerle frente está expresando esta sensación de que el sujeto no es muy hombre que digamos. Esto no tiene nada que ver con la opinión de que el sujeto merece ser culpado –si tuviera que sentir algo, no es culpa sino vergüenza– y es absolutamente compatible con la creencia de que en realidad no tuvo más opción que hacer lo que hizo. En efecto, depende de esta creencia. Sentimos cierto desprecio por esta persona porque reconocemos que no podemos esperar más de ella.

Una amenaza coercitiva suscita en su víctima un deseo –por ejemplo, el de evitar el castigo– tan poderoso que la llevará a realizar la acción exigida sin importar si quiere realizarla o si considera que sería razonable que lo hiciera. Ahora bien, un ofrecimiento también puede suscitar en la persona que lo recibe un

deseo –por ejemplo, el de obtener el beneficio–, que sea igualmente irresistible. Esto sugiere que una persona puede ser coaccionada por un ofrecimiento tanto como por una amenaza. Sin embargo, sería demasiado precipitado llegar a la conclusión de que un ofrecimiento es coercitivo siempre que su destinatario es incapaz de rechazar el beneficio que le permite obtener. Pues la única condición *necesaria* de la coacción es que la persona no tenga más opción que someterse o esté convencida de que no la tiene. En consecuencia, aunque alguien sea incapaz de soportar la fuerza motivadora del deseo de un beneficio que se le ofrece, el ofrecimiento puede no ser coercitivo.

Supongamos que alguien recibe un ofrecimiento que lo invita a llevar a cabo una acción que ya quería y tenía la intención de realizar; supongamos, además, que el beneficio que el ofrecimiento le brinda es algo que ha deseado por mucho tiempo, pero que nunca pudo obtener y que cree que sería completamente razonable tener; y supongamos también que lo que en realidad lo lleva a realizar la acción en cuestión es su deseo de ese beneficio. Este deseo puede resultarle demasiado fuerte para soportarlo. Pero no hemos supuesto solamente que él quiere llevar a cabo esta acción. También estamos suponiendo que no le molesta en absoluto ser motivado por el deseo que en definitiva lo motiva cuando la realiza. No tiene deseos ni inclinación a resistirse al deseo que lo impulsa a cumplir con los términos del ofrecimiento, y tampoco lamenta de ninguna manera sentirse motivado por él.

Sin duda, esta persona no está coaccionada, puesto que la coacción debe incluir una violación de la autonomía de su víctima. Necesariamente, la víctima de la coacción es impulsada de algún modo contra su voluntad o su voluntad se sortea de algún modo, y esta condición no se cumple en la situación que analizamos. La autonomía de la persona no se ve afectada ni en lo que hace ni en el motivo con que lo hace en esta situación a causa del ofrecimiento al que responde. El hecho de que el deseo que la induce sea irresistible concuerda con su autonomía, ya que ella se identifica por completo con este deseo. Su impulso –aunque, en efecto, está más allá de su capacidad de control– de ninguna manera la aleja de perseguir sus propias metas.

Un ofrecimiento es coercitivo, por otro lado, cuando la persona que lo recibe es inducida a cumplirlo por un deseo que no sólo es irresistible, sino que superaría si pudiera.[13] En ese caso, la persona es impulsada por un deseo por el que no desea ser impulsada. Cuando pierde el conflicto en su interior, el resultado es que es motivada contra su propia voluntad a hacer lo que hace. Así, un hombre que prefiere la fama a la oscuridad, pero que no quiere ser motivado por su preferencia puede, a pesar de todo, encontrar que es incapaz de rechazar un ofrecimiento que lo hará famoso, a pesar de todos sus esfuerzos para superar su deseo de buscar la fama. Este hombre es coaccionado a hacer cualquier cosa que haga para cumplir con los términos del ofrecimiento, sin importar si es algo que ya quería y tenía intención de hacer o no. Ello se debe a que, cuando actúa, no quiere que su voluntad sea su voluntad. Actúa siguiendo un impulso que viola sus propios deseos.[14]

Por supuesto, el carácter irresistible del deseo que suscita una *amenaza* es, en forma similar, insuficiente en sí mismo para convertir la amenaza en coercitiva. Una amenaza coercitiva, como un ofrecimiento coercitivo, sólo lo es porque también viola la autonomía de su víctima. Ahora bien, la autonomía de una persona puede ser violada por una amenaza de la misma manera en que esta violación es realizada por un ofrecimiento coercitivo. Así, el hombre que teme de manera incontrolable ser picado por una

13 Nozick niega que los ofrecimientos puedan ser coercitivos; de hecho, define la coacción en términos de amenazas, pero no toma en cuenta los tipos de consideraciones que me condujeron a interpretar algunos ofrecimientos como coercitivos. Él limita su análisis en puntos cruciales a lo que él llama *el Hombre Racional*: alguien que es "capaz de resistir las tentaciones que cree que debe resistir" (*op. cit.*, p. 460). Por supuesto, esto le impide siquiera considerar las clases de amenazas y las clases de ofrecimientos que me parecen peculiarmente coercitivos. Su uso del término *tentación*, a propósito, me parece algo impreciso. Presumiblemente, su Hombre Racional es capaz no sólo de resistir la tentación, sino también de dominar esos deseos e impulsos que –si bien no está en absoluto *tentado* a rendirse a ellos– amenazan con superar sus esfuerzos por dirigir su comportamiento únicamente según los dictados de su razón.

14 Para obtener más ejemplos de algunos de los conceptos empleados aquí y más adelante, véase "La libertad de la voluntad y el concepto de persona", en este mismo volumen. Ese ensayo desarrolla una concepción de la libertad de la voluntad en cuyos términos puede decirse que la coacción, como se la entiende aquí, priva a la víctima de su libre albedrío.

abeja puede ser amenazado con este castigo y sucumbir a su temor a pesar de todos sus esfuerzos por superarlo. En ese caso, es inducido contra su propia voluntad a someterse a la amenaza, y esto implica que es coaccionado cualquiera sea su actitud hacia la acción que realiza bajo amenaza.

Sin embargo, las amenazas irresistibles son coercitivas aun cuando no conduzcan a sus víctimas hacia esta clase de derrota interior. En efecto, *toda* amenaza efectiva cumple con la condición de que la coacción debe incluir una violación de la autonomía de su víctima. Si bien una persona es coaccionada a actuar como lo hace sólo cuando es motivada hacerlo por un deseo irresistible, actúa de alguna manera contra su propia voluntad cuando se somete a *cualquier* amenaza. Al someterse a una amenaza, una persona invariablemente hace algo que en realidad no quiere hacer. De allí que las amenazas irresistibles, a diferencia de los ofrecimientos irresistibles, sean necesariamente coercitivos.

¿Cómo debemos explicar esto? ¿Qué es lo que en una amenaza efectiva implica una violación de la autonomía de su víctima? La respuesta puede parecer obvia cuando la amenaza hace que la víctima lleve a cabo una acción que, de otra manera, habría preferido no realizar. Pero, a veces, una amenaza coincidirá con los deseos de su víctima y la inducirá en la misma dirección en la que ellos lo habrían hecho. Consideremos a un hombre que, en un marco mental ampliamente benévolo, decide salir a caminar y darle el dinero que tiene en el bolsillo a la primera persona que encuentre en la calle. La primera persona que encuentra le pone una pistola en la cabeza y lo amenaza de muerte a menos que le entregue su dinero. El hombre está aterrado, olvida su intención original en medio de su miedo y le entrega el dinero a fin de escapar de la muerte.

Aquí, la acción realizada es la que el agente quería y tenía la intención de llevar a cabo; si no hubiera sido coaccionado, la habría llevado a cabo por su cuenta. Más aun, no habría hecho frente a la amenaza aunque hubiera podido; por el contrario, como en realidad prefiere entregar su dinero que morir, sin duda habría peleado contra cualquier impulso de resistencia que pudiera haber surgido en él. Entonces, ¿en qué consiste la coacción? ¿Dónde está menoscabada la autonomía del hombre?

Es verdad que el hombre realmente prefiere ser inducido por el deseo de salvar su vida antes que por el deseo de conservar su dinero; prefiere someterse a la amenaza antes que hacerle frente. Sin embargo, éstas no son sus únicas alternativas: la sumisión y la resistencia no son las únicas respuestas a una amenaza. También es posible que una persona que recibe una amenaza no se sienta inducida por ella y que se abstenga de tomarla en cuenta. En este caso, el hombre podría haber entregado su dinero con su intención benévola original más que con la de salvar su vida. Así, habría *acatado* la amenaza, pero no habría sido coaccionado a hacerlo. Su motivo al actuar habría sido el motivo por el que quería actuar y no habría existido violación alguna de su autonomía.

Al parecer, una amenaza sólo es coercitiva, entonces, cuando el motivo por el que lleva a su víctima a actuar es un motivo por el que preferiría no actuar. Sin embargo, en realidad esta formulación de la condición no es del todo correcta. Supongamos que P amenaza con castigar a Q por hacer A, que Q todo el tiempo quería abstenerse de hacer A y tenía la intención de no hacerlo, pero que la amenaza enfurece tanto a Q que se siente irresistiblemente inducido –a pesar de sus esfuerzos por superar su ira rencorosa– a *hacer frente* a la amenaza y, por tanto, a hacer A. En este caso, la amenaza parece ser la causa de que Q actúe por un motivo por el que preferiría no sentirse inducido. Sin embargo, mientras que su autonomía es violada, sin duda, por la furia que lo embarga, ciertamente no es coaccionado por P a hacer A. Es evidente que una amenaza sólo es coercitiva cuando hace que su víctima lleve a cabo, por un motivo por el que preferiría no ser inducida, una acción que *acata* la amenaza.

Ahora bien, ¿por qué es invariablemente verdad que la persona que se somete a una amenaza lo hace por un motivo por el que preferiría no ser inducida?[15] No es una respuesta adecuada, o por

15 Nozick destaca (*op. cit.*, pp. 46 y ss.) que en la naturaleza de las amenazas está la noción de que una persona no pensará nunca que es amenazada por su bien (explica razonablemente cómo responder a los contraejemplos obvios de esto). Sin embargo, el uso que se le puede dar a este punto no me resulta del todo claro. El solo hecho de que una persona esté en una situación en la que preferiría no estar ciertamente no implica, en sí mismo, que no sea completamente autónoma

lo menos no es adecuadamente precisa, decir que, en esos casos, la persona siempre es inducida por el miedo: esa persona puede tener su propia motivación al actuar para obtener un beneficio, y es difícil especificar las diferencias que puede haber entre el deseo de obtener un beneficio y el temor de perderlo. Una respuesta un poco mejor es que una persona que se somete a una amenaza (y que no sólo la acata) necesariamente lo hace a fin de evitar un castigo. Es decir, su motivación no es mejorar su condición, sino impedir que empeore. Esto parece suficiente para explicar el hecho de que preferiría tener un motivo diferente para actuar.[16] También sugiere por qué en principio nos pronunciamos contra el hecho de amenazar a la gente y por qué, en general, se cree que las amenazas –a diferencia de los ofrecimientos– requieren justificación. Alguien que recibe una amenaza no tiene nada para ganar de ella y todo para perder. A diferencia de un ofrecimiento, una amenaza expone a una persona al riesgo de un castigo adicional sin proporcionarle ninguna oportunidad de obtener un beneficio que, de otra manera, no habría estado disponible para ella.

Cuando P ejerce coacción sobre Q para que haga A, entonces Q no hace A libremente o guiado por su libre albedrío. También es verdad que, en cierto sentido, P somete a Q a su voluntad, o que reemplaza la voluntad de Q con la propia: el motivo que tiene Q no es el que Q desea, sino el que P le hace tener. Ahora bien, no hablamos de coacción excepto cuando una persona impone su voluntad de esta manera sobre otra. Tenemos buenas razones para notar especialmente los papeles que desempeñan en nuestras vidas

---

en cualquier elección que haga entre las alternativas que la situación le proporciona. Ello se debe a que una persona puede hacer frente a una amenaza o, del modo descripto anteriormente, acatarla sin someterse a ella; y en esos casos su autonomía no está en absoluto dañada, a pesar de que no habría elegido ser amenazada. El punto central del análisis de Nozick se desdibuja un poco, creo, por el ejemplo que él plantea. No es para nada un ejemplo de amenaza, sino que se refiere, en cambio, a alguien que se ha roto la pierna y que está eligiendo entre un yeso decorado y uno sin decorar. Esta elección es diferente de todas las elecciones que suelen tener las personas que han sido amenazadas, pues ellas pueden actuar como si la amenaza no hubiera existido.

16 Para consultar un análisis sugestivo de este y otros puntos relacionados, véase Gerald Dworkin, "Acting Freely", en *Nous*, IV, 1970, pp. 367-383.

las acciones de otros hombres y para distinguirlos de los papeles que desempeñan circunstancias de otra índole. Nuestros modos de sobrellevar y regular estas dos clases de condiciones son muy diferentes. Pero el efecto de la coacción sobre su víctima, en virtud de la cual la autonomía o la libertad de la víctima se ven debilitadas, no se debe, en esencia, al hecho de que la última esté sometida a la voluntad de otra persona.

Consideremos las siguientes situaciones. Supongamos, en primer lugar, que un hombre llega a una bifurcación en el camino, que alguien ubicado en una colina junto al lado izquierdo de la bifurcación lo amenaza con causar una avalancha que lo aplastará si toma ese camino, y que el hombre sigue hacia la derecha a fin de satisfacer el deseo dominante de preservar su propia vida. Luego, supongamos que cuando el hombre llega a la bifurcación no encuentra a nadie que lo amenace, sino que nota que, a causa de la condición natural de las cosas, será aplastado por una avalancha si toma el camino de la izquierda, y que es inducido en forma irresistible, por su deseo de vivir, a seguir por el camino de la derecha. Hay diferencias interesantes entre estas situaciones, ciertamente, pero no existe un fundamento para considerar que el hombre esté guiado por su libre albedrío o actuando con mayor o menor libertad en un caso que en el otro. Si es moralmente responsable de su decisión o su acción en cada caso no depende de la fuente del daño que está tratando de evitar, sino de la forma en la que su deseo de evitarlo opera dentro de él.[17]

Por supuesto, solemos enojarnos más cuando otra persona coloca obstáculos en nuestro camino que cuando lo hace nuestro entorno. Sin embargo, lo que explica este mayor enojo no es el amor

17 Se podría esgrimir el argumento de que siempre es deseable, hasta cierto punto, hacer frente a una amenaza, sin importar sus términos, y que *prima facie* no existe una conveniencia correspondiente para hacer frente al entorno natural. Si argumentos como éste fueran sólidos, sería más difícil justificar una acción realizada con el fin de evitar un castigo derivado de una amenaza que justificar una acción realizada a fin de escapar de un daño comparable provocado por el entorno. Sin embargo, esto no significaría que la responsabilidad moral de una persona por lo que hace esté afectada en forma diferente si sus motivos le fueron impuestos por otra persona o si surgen de su encuentro con condiciones naturales en las que no intervino otra persona.

a la libertad. Es el orgullo; o algo que está más cerca del orgullo, un sentido de injusticia. Sólo otra persona puede *ejercer coacción* sobre nosotros, o interferir con nuestra libertad *social* o *política*, pero esto no es más que una cuestión de terminología útil. Cuando una persona elige actuar a fin de obtener un beneficio o a fin de escapar de un daño, la medida en que su elección es autónoma y la medida en que actúa libremente no dependen del origen de las condiciones que la llevan a elegir y a actuar como lo hace. La voluntad de un hombre puede no pertenecerle aun cuando no sea inducido por la voluntad de otro.[18]

18 Este ensayo tendría menos errores si yo hubiera podido atender mejor a los valiosos comentarios que, sobre una versión anterior, me hicieron Peter Hacker, Anthony Kenny, Sydney Morgenbesser y Joseph Raz.

4
Tres conceptos de *acción libre*

I

En muchas situaciones, una persona lleva a cabo una acción porque la prefiere a cualquier otra de las que piensa que tiene disponibles o porque se siente atraída hacia ella con mayor fuerza que hacia cualquier otra y, sin embargo, se resiste a afirmar sin reservas que ha actuado voluntariamente. Es posible que reconozca que hizo lo que en *cierto* sentido quería hacer y que comprendía bastante bien lo que quería y lo que hizo. No obstante, al mismo tiempo puede pensar que es pertinente y justificable disociarse en cierta forma de su acción, quizá diciendo que lo que hizo no era algo que *realmente* quería hacer o que no era algo que realmente *quería* hacer. Estas situaciones se clasifican en varios tipos distintos.

En las situaciones del Tipo A, la sensación de una persona de que actuó involuntariamente deriva del hecho de que las circunstancias externas en las cuales actuó no eran, según su percepción, acordes con sus deseos. Desde luego, una persona casi siempre tiene la posibilidad de imaginar una situación que le agradaría más que aquella en la que en realidad está. Sin embargo, existe una diferencia sustancial –a menudo, bastante fácil de discernir, aunque difícil de explicar con precisión– entre reconocer que una situación dista mucho de ser ideal y sentirse descontento con ella o resistirse a ella en forma activa. La discordancia entre la realidad y el deseo que caracteriza las situaciones del Tipo A da lugar a esto último, y no sólo a lo primero: no se trata únicamente de que haya otra situación imaginable en la que el agente preferiría estar, sino

que lamenta o está enojado con la situación con la que, de hecho, tiene que lidiar.

Supongamos que la razón por la que alguien llevó a cabo cierta acción fue que consideró que era el mal menor entre los que tenía que elegir. Dadas las alternativas que enfrentaba, prefirió sin reservas la que llevó adelante. Esto proporciona la justificación para afirmar que hizo lo que quería hacer. Sin embargo, las alternativas que enfrentaba constituían un conjunto del cual él no quería tener que elegir; estaba disgustado con la necesidad de tener que hacer esa elección. Esta discrepancia, entre el mundo como era y como él quería que fuera, es la que respalda su afirmación de que no actuó en forma totalmente voluntaria.

En las situaciones del Tipo B, las circunstancias internas de su acción no concuerdan con los deseos del agente. Lo que motiva su acción es un deseo por el cual, dadas las alternativas que enfrenta, no quiere ser inducido a actuar. Existe un conflicto en su interior, entre un deseo de primer orden de hacer lo que en realidad hace y una volición de segundo orden de que este deseo de primer orden no sea efectivo para determinar su acción. En otras palabras, quiere sentirse motivado en forma efectiva, respecto de las alternativas que enfrenta, por algún otro deseo, que no es el que realmente lo induce a actuar como lo hace.

El hecho de que esta persona niegue haber actuado en forma totalmente voluntaria refleja su sensación de que, en el conflicto del cual emergió su acción, fue vencida por una fuerza con la cual, aunque surgió de su interior, no se identificaba. Por ejemplo, puede haber luchado en vano contra un fuerte anhelo (deseo efectivo de primer orden) ante el cual no quería sucumbir (volición de segundo orden, derrotada). Entonces, su intento de disociarse de lo que hizo expresa su interpretación de que ha sido impotente ante un deseo que la condujo, contra su voluntad, a hacer lo que hizo, a pesar de que ella prefería otra acción.

La desdicha mundana y el conflicto interno, como los que aquí se tratan, influyen en forma diferente en las responsabilidades morales de los agentes que los sufren. En virtud de la discrepancia entre el deseo que motiva su acción y el deseo por el cual quiere ser motivado, el agente en una situación del Tipo B puede no ser mo-

ralmente responsable de lo que hace. El deseo que lo induce sin dudas es, en cierta manera, irrefutablemente suyo. Sin embargo, lo induce a actuar contra su propia voluntad o contra la voluntad que quiere tener. En este sentido, le es ajeno, lo que puede justificar que se considere que fue inducido a hacer lo que hizo, en forma pasiva, por una fuerza de la que no se lo puede juzgar moralmente responsable.

Por otra parte, el hecho de que alguien enfrente alternativas entre las cuales no quiere tener que elegir carece de efectos sobre su responsabilidad moral por la acción que elige llevar a cabo. No hay razón por la cual no debería adjudicarse mérito o culpa a una persona –en otras palabras, por la que una persona no debería ser considerada moralmente responsable– por el modo en que actúa en situaciones que preferiría haber evitado así como por su comportamiento en situaciones en las que le agrada estar. Así, el hecho de que una persona actúe en una situación del Tipo A no proporciona ningún tipo de fundamento para negar que sea moralmente responsable de lo que hace.

Entre las situaciones del Tipo A, hay muchas en las que el agente es amenazado con un castigo, que le será impuesto por otra persona o por fuerzas impersonales, a menos que realice cierta acción. Ahora bien, si suponemos que verse *coaccionado* a hacer algo excluye el ser moralmente responsable de hacerlo, una amenaza no es coercitiva cuando la persona amenazada cree con razón que puede hacerle frente si elige hacerlo. Esto se debe a que, en ese caso, si se somete a la amenaza realizará una acción respecto de la cual cree tener alternativa; la realiza, entonces, porque ella misma decide hacerlo. Y su acción es, por consiguiente, meritoria o censurable –es decir, la persona *es* moralmente responsable de realizarla– según si realizarla o no es moralmente preferible a la resistencia, dadas las circunstancias en que elige llevarla a cabo.

Las amenazas coercitivas, por otra parte, implican castigos que el destinatario de la amenaza no puede elegir sufrir en forma efectiva. Su inclinación a evitar la consecuencia indeseable que enfrenta es irresistible; es imposible que logre aceptar esa consecuencia. Cuando el carácter irresistible de esta inclinación o deseo es lo que explica la acción que realiza, el destinatario de una amenaza

no es moralmente responsable de lo que hace, no más que alguien que realiza una acción que puede explicarse por una compulsión irresistible que se origina en su interior.

La situación de una persona que sucumbe a una amenaza porque es incapaz de enfrentarla no pertenece, por consiguiente, al Tipo A. Tampoco es necesariamente del Tipo B, ya que no hay razones para suponer que una persona que actúa debido al carácter irresistible de un deseo preferiría, dadas sus alternativas, ser motivada por un deseo diferente. La situación es, por tanto, de un tipo distinguible C, cuya característica especial es que el agente actúa debido al carácter irresistible de un deseo sin que intente impedir que ese deseo determine su acción. (El agente que se encuentra en una situación del Tipo C es, en ciertos aspectos, análogo a lo que en otro ensayo denominé *agente inconsciente*, que, en parte, puede explicar la repugnancia de la coacción.)* Él no es vencido por el deseo, como en las situaciones del Tipo B, ya que no le opone una volición de segundo orden. Tampoco es autónomo dentro de los límites de un conjunto de alternativas insatisfactorias, como en las situaciones del Tipo A, ya que su acción no es el resultado de una elección efectiva en cuanto a qué hacer.

A la luz de algunas de las observaciones de Don Locke,[1] es necesario aclarar especialmente dos puntos en relación con las situaciones del Tipo C. En primer lugar, una persona puede actuar para satisfacer un deseo que no puede, de hecho, resistir, y, sin embargo, es posible que el carácter irresistible del deseo no sea lo que explica su acción. Por ejemplo, ella puede no ser consciente de que el deseo es irresistible y realizar –por razones que no están relacionadas con su incapacidad de resistirlo– la misma acción hacia la cual habría sido conducida si las condiciones fueran otras. En ese caso, su situación no es del Tipo C. En segundo lugar, el hecho de que alguien actúe debido al carácter irresistible de un deseo no significa que actúa con pánico o con un impulso repentino. Una persona

* Véase "La libertad de la voluntad y el concepto de persona", capítulo 2 de este volumen. [N. de T.]

1 Don Locke, "Three concepts of free action: I," en *Proceedings of the aristotelian society*, vol. comp. XLIX, 1975, pp. 95-112. Este ensayo fue escrito como respuesta al de Locke.

puede creer que no es capaz de resistir cierto deseo y, por tanto, proceder a satisfacerlo con calma resignación, sin experimentar la incontrolable fuerza compulsiva que en realidad sentiría si intentara no sastisfacerlo.

Tanto en las situaciones del Tipo B como en las del Tipo C, el agente es inducido a actuar sin el acuerdo de una volición de segundo orden: en el primer caso, porque su volición de segundo orden es derrotada y, en el último, porque no hay una volición de segundo orden que desempeñe una función en la economía de sus deseos. Por otra parte, el agente que se encuentra en una situación del Tipo A aprueba el deseo que lo induce a actuar. A pesar de su insatisfacción con una situación en la que encuentra que el deseo merece su respaldo, se siente satisfecho, dada la situación, de ser inducido por el deseo. Por tanto, su acción concuerda con una volición de segundo orden, y la renuencia con la que actúa es de carácter diferente de la renuencia con la que actúan los agentes en las situaciones de los Tipos B y C. Esta diferencia se refleja en el hecho de que él puede ser moralmente responsable de lo que hace, mientras que aquéllos no lo son.

Consideremos si es posible identificar la clase de acciones libres con la clase de acciones –llamémosla *W*– que *no* se llevan a cabo en las situaciones del Tipo B o del Tipo C, suponiendo que todas las acciones de esta clase se realizan con un entendimiento y una intención adecuados. Notemos que *W* se parece a la clase de acciones voluntarias definidas por la noción de voluntad de Locke. Sin embargo, *W* es más amplia que esa clase: incluye no sólo a todos los miembros de esta última, sino también las acciones realizadas en las situaciones del Tipo A.

II

Debido a esta diferencia entre ambas clases, la afirmación de Locke de que la disposición voluntaria no puede ser una condición *necesaria* para la acción libre no tiene fuerza contra la opinión de que las acciones libres deben pertenecer a *W*. A fin de respaldar su afir-

mación, Locke menciona el ejemplo de una persona que actúa para hacer lo que considera su deber, pero que desea que la acción que lleva a cabo no le fuera exigida desde un punto de vista moral. Locke observa que si bien esta persona actúa contra su voluntad, "sería poco razonable sugerir que, por ese motivo, no está actuando en forma libre, y menos que carece de responsabilidad por actuar como lo hace". Sin embargo, dado que la situación del moralista reacio es del Tipo A, y no del Tipo B o del Tipo C, su acción pertenece a *W*. Así, la observación de Locke de que actúa tanto en forma libre como responsable no crea dificultades para la opinión de que pertenecer a *W* es una condición necesaria para la acción libre.

Locke también sostiene que la disposición voluntaria no puede ser una condición *suficiente* para la acción libre. El fundamento que brinda para esta afirmación no deja de ser en absoluto pertinente debido a la diferencia entre su clase de acciones voluntarias y *W*. Locke argumenta que si se considera que la disposición voluntaria es suficiente para que una acción sea libre –también se puede decir lo mismo si se considera que pertenecer a *W* es suficiente–, un adicto por voluntad propia actúa libremente cuando consume la droga a la que es adicto, mientras que no sucede lo mismo con un adicto contra su voluntad. Ésta es la conclusión que se desprende, y Locke afirma que es "poco razonable, dado que ambos son adictos".

Sin embargo, ¿dónde reside la falta de razonabilidad? El hecho de que ambos sean adictos sólo significa que ninguno puede abstenerse de consumir la droga o, si se quiere, que ninguno es libre de abstenerse de hacerlo. Esto no resuelve la cuestión, puesto que dista mucho de ser evidente que una persona que no es libre de abstenerse de realizar cierta acción no pueda ser libre para realizarla, o que no pueda realizarla libremente. Después de todo, ¿por qué debería interpretarse que la libertad de una persona respecto de la realización de una acción tiene algo que ver, en esencia, con su libertad respecto de la realización de otra? Sin embargo, es evidente que Locke da por sentado que hacer algo libremente implica poder abstenerse de hacerlo o evitarlo. Más avanzado su ensayo, adopta esta suposición más abiertamente y la convierte en un punto central de su propia explicación de la

acción libre. No obstante, en ninguna parte de su ensayo ofrece argumentos para apoyarla.

En mi opinión es muy poco razonable sostener, como también hace Locke, que los dos adictos tienen la misma responsabilidad moral por consumir la droga. Creo que es decisivo, al respecto, que la adicción del adicto por voluntad propia puede no desempeñar ningún papel en la explicación de su acción. En efecto, lo que explica que consuma la droga puede ser, en especial si no es consciente de su adicción, exactamente lo mismo que lo que explica el consumo de una droga por parte de alguien que no es adicto y que sólo la consume porque, con franqueza, le agrada hacerlo. En ese caso, con seguridad, la responsabilidad moral por su acción del adicto por voluntad propia se corresponde con la del que no es adicto y consume la droga, y no con la del adicto que la consume contra su voluntad y cuya acción puede explicarse sólo en términos de su adicción.

Lo que la acción del adicto por voluntad propia revela acerca de él es lo mismo que revela la acción del que no es adicto. No es lo mismo que lo que revela la acción del adicto contra su voluntad. Me parece, entonces, que el significado moral de la acción del adicto por voluntad propia es el mismo que el de la acción del que no es adicto y diferente del significado moral de lo que hace el adicto contra su voluntad. De paso, notemos que las acciones del adicto por voluntad propia y las del que no es adicto pertenecen a *W*, mientras que las del adicto contra su voluntad, no.

Al evaluar la respectiva responsabilidad moral del adicto por voluntad propia y el que lo es contra su voluntad, Locke tiende a hacer caso omiso de la distinción entre realizar una acción que uno no puede evitar realizar y realizar una acción *porque* uno no puede evitar realizarla. Aparentemente, esto es lo que hace posible que él considere equivocadamente que el problema de evaluar la responsabilidad moral en el caso de los dos adictos es realmente análogo al problema de evaluarla en el caso de los dos pilotos secuestrados. Su comentario acerca de los pilotos –que sería extraño atribuir al piloto dispuesto a desviar su curso cierta responsabilidad por dirigirse a Cuba, que no se le puede atribuir de igual manera al piloto que lo hace contra su voluntad– no respalda, como él cree, su afirmación de que la responsabilidad de ambos adictos es la misma.

Los dos pilotos, como los describe Locke, actúan por la misma razón –para evitar ser asesinados–, aunque a uno de ellos le agrada tener esta razón para volar a Cuba, mientras que no sucede lo mismo con el otro. Ahora bien, precisamente debido a que actúan por la misma razón o porque sus acciones se explican de la misma manera, los pilotos son igualmente responsables de lo que hacen. Desde luego, en este aspecto crucial, su caso difiere del caso de los dos adictos, ya que lo que explica la acción de un adicto no explica la acción del otro. Si supusiéramos que el piloto dispuesto a desviar su curso se dirigió a Cuba porque quería ver a su amante en La Habana, y que el deseo de evitar ser asesinado no fue lo que en realidad lo motivó, el caso de los dos pilotos *sería* más parecido al de los dos adictos. Sin embargo, creo que también juzgaríamos que la responsabilidad del piloto dispuesto a desviar su curso difiere de la responsabilidad del piloto que lo hace contra su voluntad.

Los análisis de Locke sobre los adictos y los pilotos no brindan ningún fundamento que permita negarse a identificar a *W* con la clase de acciones libres ni para aceptar su afirmación de que las acciones libres deben poder evitarse. Es claro que el adicto contra su voluntad no actúa de manera libre ni responsable cuando consume la droga, y que el adicto por voluntad propia no es más capaz de evitar consumirla que él. No obstante, los dos adictos difieren en lo que los conduce a la acción. En virtud de esa diferencia, sus responsabilidades morales por aquellas acciones difieren. La opinión de que no hay diferencia en la libertad con la que actúan no es respaldada, entonces, por las consideraciones relacionadas con la responsabilidad moral.

III

Parece que las acciones que pertenecen a *W* podrían ser realizadas por un individuo cuya vida mental y su comportamiento físico estuvieran determinados por algún tipo de manipulación de su condición fisiológica efectuada por el Demonio/neurólogo de Locke. Por tanto, identificar a *W* con la clase de acciones libres significa-

ría, en apariencia, admitir que el D/n podría asegurar que su sujeto actuara libremente. ¿Se puede objetar esto?

Distingamos dos situaciones, en esencia diferentes, haciendo caso omiso de los difíciles problemas que plantean los posibles casos mixtos y fronterizos. En la primera situación, el D/n manipula a su sujeto en forma continua, como una marioneta, de manera que cada uno de los estados mentales y físicos del sujeto es el resultado de la intervención específica del D/n. En ese caso, el sujeto no es, en absoluto, una persona. Su historia es totalmente episódica y carece de cohesión interna. Los temas identificables que pueda revelar no están arraigados internamente; no se puede interpretar que constituyen la propia naturaleza del sujeto o que pertenecen a ella. En cambio, son proporcionados en forma gratuita por un agente externo al sujeto. Por cierto, los estados de ánimo instantáneos del sujeto pueden ser tan ricos como los de una persona; pueden incluir deseos y voliciones de segundo orden o tener, incluso, estructuras más complejas. No obstante, el sujeto no tiene carácter o temperamento propios, y no hay razones para esperar de él –salvo por derivación, siempre que haya razones para esperarlo del D/n– ni siquiera la mínima continuidad e inteligibilidad esenciales relativas a ser una persona. Creo que hay instancias de su conducta que pueden excluirse, razonablemente, de *W* sobre la base de que, debido a que el sujeto carece de toda autonomía, ellas no pueden ser consideradas legítimamente sus acciones.

La otra posibilidad es que el D/n proporcione a su sujeto un carácter o programa estable, que luego no altere –o al menos no con demasiada frecuencia– y que las respuestas mentales y físicas posteriores del sujeto a sus ámbitos externo e interno estén determinadas por este programa y no por otra intervención del D/n. En ese caso, no hay razones para negar que hay instancias de la conducta del sujeto que podrían pertenecer a *W*. Desde mi punto de vista, tampoco hay razones de peso para oponerse a admitir que el sujeto puede actuar libremente ni para considerar que es capaz de ser moralmente responsable de lo que hace.

Puede tornarse moralmente responsable, suponiendo que esté programado en forma adecuada, de la misma manera que sucede con los demás: identificándose con algunos de sus propios deseos

de segundo orden, de manera que no sólo sean deseos que simplemente tiene o encuentra en su interior, sino deseos que adopta o defiende. En virtud de la identificación de una persona con uno de sus propios deseos de segundo orden, ese deseo se convierte en una volición de segundo orden. Y, así, la persona *asume* la responsabilidad de los deseos pertinentes de primero y segundo orden y de las acciones hacia las que estos deseos la conducen.

Locke sugiere que no se puede considerar que el sujeto de un D/n actúa libremente, porque lo que hace o los deseos y las voliciones de segundo orden que tiene no dependen del sujeto ni están bajo su control. Es evidente que él interpreta las nociones de *depende de X* y *bajo el control de X* en el sentido de que nada depende de una persona ni está bajo su control si un tercero determina su ocurrencia. Pienso que estas nociones deben interpretarse de manera diferente, al menos en este contexto. Lo que está en juego en su aplicación no es tanto asunto de los orígenes causales de las situaciones en cuestión, sino la actividad o la pasividad de *X* respecto de esas situaciones.

Ahora bien, una persona es activa respecto de sus propios deseos cuando se identifica con ellos, y es activa respecto de lo que hace cuando lo que hace es el resultado de su identificación con el deseo que la induce a hacerlo. Sin esa identificación, la persona es un observador pasivo de sus deseos y de lo que hace, independientemente de si las causas de sus deseos y de lo que hace son el trabajo de otro agente, de fuerzas externas impersonales o de procesos internos de su propio cuerpo. En cambio, es imposible que una persona sea un observador pasivo de sus voliciones de segundo orden. Ellas *constituyen* su actividad –es decir, el hecho de que sea activa y no pasiva– y no puede plantearse la cuestión de si la persona se identifica con ellas o no. No tiene sentido preguntar si alguien se identifica con la identificación que tiene de sí mismo, a menos que la intención sea simplemente preguntar si su identificación es sincera o total.

Hay que reconocer que esta noción de identificación es algo desconcertante, y no estoy seguro de cómo explicarla. No obstante, creo que capta algo bastante fundamental de nuestras vidas internas, y merece un papel central en la fenomenología y en la filosofía

de la mentalidad humana. En vez de intentar proporcionar el análisis que la noción requiere, me limitaré a hacer una declaración: en la medida en que una persona se identifica con los resortes de sus acciones, asume la responsabilidad por esas acciones y adquiere una responsabilidad moral respecto de ellas; más aun, la cuestión de cuál es la causa de las acciones y su identificación con sus resortes no tiene relación con la cuestión de si lleva a cabo las acciones de manera libre o si es moralmente responsable de realizarlas.

El hecho de que el D/n haga que su sujeto tenga ciertos deseos de segundo orden y se identifique con ellos no afecta, entonces, el sentido moral de la adquisición, por parte del sujeto, de las voliciones de segundo orden con las que, de ese modo, es dotado. No hay paradoja en la suposición de que un D/n podría crear un agente moralmente libre. Por cierto, podría ser razonable sostener que el D/n es demasiado responsable desde un punto de vista moral de lo que hace su sujeto libre, al menos en la medida en que se lo pueda considerar claramente responsable de anticipar las acciones del sujeto. Sin embargo, esto no implica que no pueda atribuírsele también al sujeto la responsabilidad moral total por aquellas acciones. Es muy posible que más de una persona tenga total responsabilidad moral por el mismo acontecimiento o acción.

IV

Una acción puede pertenecer a *W* incluso cuando el agente la realiza bajo presión, es decir, bajo amenaza, pero cuando la sumisión del agente no es el resultado del carácter irresistible de su deseo de evitar el castigo con el que es amenazado. Si se identifica a *W* con la clase de acciones libres, de ello se desprende, por supuesto, que las acciones pueden llevarse a cabo libremente, incluso cuando se las realiza bajo presión. Esto suena algo extraño. No obstante, debe notarse que, por cierto, la consecuencia no sería inaceptable para Locke.

Él está dispuesto a admitir que un empleado bancario actúa libremente cuando, para evitar que le disparen, accede a los pedi-

dos de un asaltante armado. Y explica que se puede decir que el empleado actuó libremente, porque "es su decisión y sólo su decisión si entregar los fondos o arriesgar su vida, y lo que haga dependerá de él, de sus deseos y de sus preferencias, de nadie más, ni siquiera del asaltante". Aparentemente, Locke no supone que el empleado es empujado por su deseo de evitar que le disparen, o que su acción deriva de la creencia de que no puede evitar sucumbir a este deseo. En otras palabras, interpreta que la situación del empleado es del Tipo A y que su acción pertenece a *W*: el empleado actúa bajo presión, pero no es coaccionado.

En esta suposición, el hecho de afirmar que el empleado actúa libremente –aunque no se pueda decir que actúa de acuerdo con su voluntad– se debe a que la acción del empleado no es el resultado de haber sido forzado, en forma inevitable, a actuar como lo hace. Su acción es el producto de una decisión para la cual tiene, o cree tener, una alternativa posible. Por tanto, expresa una elección genuina de su parte. Es claramente irrelevante para la autenticidad de esta elección si el empleado tiene razón o no en creer que podía haberse abstenido de entregar los fondos o podía haberlo evitado, si prefería esa alternativa. En la medida en que la libertad de la acción del empleado se basa en una elección auténtica de su parte, como sugiere Locke, parecería que la libertad de acción no requiere que ésta sea evitable.

Aunque Locke considera que el empleado actuó con libertad, no cree que sea moralmente responsable de su acción. De hecho, según Locke, es un error suponer que existe una relación necesaria entre los conceptos de acción libre y de responsabilidad moral. No obstante, las razones que ofrece para sostener esta opinión no son convincentes.

A fin de respaldar la postura de que actuar libremente no es una condición necesaria de la responsabilidad moral, Locke afirma que una persona que conduce a alta velocidad puede ser moralmente responsable de atropellar a un niño que sale corriendo de detrás de un automóvil estacionado, aunque el conductor no sea libre al atropellarlo. Pero esta afirmación parece incorrecta. Si no intervienen otros factores, el conductor no es más culpable que si hubiera evitado el accidente; y es mucho menos culpable que alguien que

atropella a un niño libremente. Estas consideraciones indican que la persona que conduce a alta velocidad no es en absoluto moralmente responsable de atropellar al niño. Sí es responsable de, por ejemplo, conducir en forma imprudente, y no hay razones para dudar de que lo hace libremente. Entonces, el ejemplo no debilita la posición de que actuar libremente es una condición necesaria de la responsabilidad moral.

¿Qué se puede decir acerca del empleado bancario? Es verdad que probablemente no lo culparíamos por acceder a las demandas del asaltante armado. Por cierto, Locke supone que ello se debe a que no lo creemos moralmente responsable de su acción, y que si la acción libre estuviera ligada con la responsabilidad moral, tendríamos que decir –incorrectamente, según él– que el empleado no actúa libremente. Pero si consideramos que el empleado actúa en forma libre, la razón por la que nos abstenemos de culparlo no es porque pensemos que no tiene responsabilidad moral de haber accedido al pedido del asaltante. Es porque juzgamos que actúa en forma razonable cuando entrega el dinero del banco en vez de su propia vida y, por tanto, no encontramos nada de qué culparlo. Así, el ejemplo no demuestra, como se supone que debería hacer, que una persona puede actuar en forma libre sin ser moralmente responsable de lo que hace.

V

Dadas las funciones que desempeñan la libertad y otros conceptos afines en nuestro esquema conceptual general, y las reflexiones acerca de estas funciones en la tradición lingüística, no dejaría de ser razonable exigir que el concepto de acción libre se interpretara de manera tal que una persona no tenga responsabilidad moral salvo por lo que hizo libremente. Sobre la misma base, también sería razonable exigir que no se interpretara que una acción se llevó a cabo libremente si fue realizada bajo presión o bajo una presión con cierto grado de severidad. El análisis de Locke, según el cual las acciones libres son aquellas que pueden evitarse, no satisface nin-

guno de estos requerimientos. Ser capaz de evitar llevar a cabo una acción no es incompatible con realizar esa acción bajo presión, ni tampoco es una condición necesaria de ser moralmente responsable de llevarla a cabo.

Ello no significa que la posibilidad de evitar una acción no tenga relación con la libertad. Sólo significa que una persona puede actuar libremente cuando no es libre de actuar de otra manera. Es decir que del hecho de que *X* hizo *A* libremente no se desprende que *X* era libre de abstenerse de hacer *A*. Así, una persona puede ser libre para hacer lo que realmente quiere hacer, y hacerlo libremente, sin gozar de libertad en el sentido de encontrarse en la posición de hacer lo que podría querer y, a la vez, ser inherentemente capaz de hacer. No me parece objetable decir que un prisionero puede llegar a actuar más libremente –es decir, llevar a cabo más acciones libremente– después de aprender las lecciones de Epicteto. Pero naturalmente seguirá habiendo tantas cosas como antes que no es libre de hacer y que sería capaz de hacer si no estuviera preso. Por tanto, no sería correcto decir que ha sido liberado, o que ha escapado a las limitaciones de su prisión, mediante el estudio de la filosofía estoica.

Identificar *W* con la clase de acciones libres, ya que permite que las acciones se lleven a cabo tanto libremente como bajo presión, no satisface uno de los requerimientos propuestos. Sí parece satisfacer el otro: pertenecer a *W* es una condición necesaria para una acción moralmente responsable. De hecho, no es posible satisfacer ambos requerimientos a la vez, porque una persona puede ser moralmente responsable de lo que hace bajo presión. Frases como "Lo hizo libremente" se emplean, en realidad, en forma algo equívoca: a veces, connotan que el agente hizo lo que hizo por propia voluntad, y, a veces, connotan su responsabilidad moral por haberlo hecho. Si hemos de tener un uso filosófico establecido y unívoco de *acción libre*, debemos decidir si es preferible satisfacer un requerimiento o el otro. Por lo que veo, no hay mucha opción entre estas alternativas.

5
Identificación y externalidad

I

Uno de los problemas centrales y más difíciles de la teoría de la acción suele formularse de un modo similar al siguiente: ¿cómo se explica la diferencia entre la clase de cosa que sucede cuando una persona levanta un brazo (digamos, para hacer una seña) y la clase de cosa que sucede cuando el brazo de una persona se levanta (digamos, por un espasmo muscular) sin que ésta lo mueva? La pregunta evoca un contraste entre acontecimientos que son acciones, en los que entran en juego las facultades superiores de los seres humanos, y aquellos movimientos del cuerpo de una persona –instancias de comportamiento que no son acciones, o simples sucesos corporales– que no hace por sí misma. Este contraste, evidentemente, puede generalizarse. Es un caso especial del contraste entre la actividad y la pasividad, que tiene un alcance considerablemente mayor.

Las acciones son instancias de actividad, aunque no son las únicas, ni siquiera en la vida humana. Tamborilear los dedos sobre la mesa, ociosa y distraídamente, con toda certeza no es un caso de pasividad: los movimientos en cuestión no ocurren sin que uno los realice. Sin embargo, tampoco es una instancia de acción, sino sólo de estar activo. A veces se pasa por alto el hecho de que hay acontecimientos en la vida humana que no son acciones ni simples sucesos, pero esto no debería sorprendernos.[1] El contraste entre ac-

1 Un resultado de pasar por alto los acontecimientos de esta clase es una exageración de la peculiaridad de lo que hacen los humanos. Otro resultado,

tividad y pasividad es fácilmente perceptible en niveles de la existencia en los que tendemos a suponer que no hay acciones. Así, una araña es pasiva con respecto a los movimientos de sus patas cuando éstas se mueven porque la araña ha recibido una descarga eléctrica. Por otro lado, la araña es activa con respecto a los movimientos de sus patas –aunque no lleve a cabo ninguna acción– cuando las mueve al desplazarse por el suelo. No debería parecernos poco natural que seamos capaces, sin caer en una mera pasividad, de comportarnos tan mecánicamente como la araña.

No es nada fácil explicar la diferencia ente estar activo y estar pasivo, y, de hecho, durante algún tiempo los filósofos han tendido a descuidar esta tarea. Aristóteles dio un paso adelante al dividir los acontecimientos de la historia de una cosa en aquellos cuyo principio motor está dentro de la cosa y aquellos cuyo principio motor está afuera. Esto es sugerente: una cosa es activa con respecto a los acontecimientos cuyo principio motor está en su interior, y pasiva con respecto a los acontecimientos cuyo principio motor está fuera de ella. Sin embargo, la distinción interior-exterior, que parece subyacer a la que se plantea entre la actividad y la pasividad, desafortunadamente no es menos difícil de entender. Está claro que los términos *adentro* y *afuera* no pueden tomarse en sus significados literalmente espaciales. Si un hombre es arrastrado por el viento, lo que lo mueve está fuera de su cuerpo. Si su cuerpo o alguna parte de él es movido por un espasmo que ocurre en sus músculos, es igualmente pasivo con respecto a ese acontecimiento aunque su principio motor esté dentro de su cuerpo. Entonces, en este caso, ¿con respecto a qué es "externo" el principio motor?

No trataré de explorar el fundamento de la distinción entre la actividad y la pasividad. En su lugar, ampliaré el alcance de la distinción con la que empecé. El contraste entre los movimientos del cuerpo de una persona que son meros sucesos en su historia, y los que son sus propias actividades, nos aleja no sólo de la vida humana hacia los reinos inferiores de la creación, sino que también

---

relacionado con el primero, es la creencia errónea de que una bipartición de los acontecimientos humanos en acciones y simples sucesos proporciona una clasificación que satisface los intereses de la teoría de la acción.

nos conduce, en virtud de la distinción análoga que puede establecerse en la esfera psicológica, al centro de la experiencia que tenemos de nosotros mismos.

En nuestros procesos intelectuales, podemos ser activos o pasivos. Concentrar la mente en algo determinado o deliberar sistemáticamente acerca de un problema son actividades que realiza una persona. Sin embargo, somos simples observadores pasivos con respecto a algunos de los pensamientos que ocurren en nuestra mente, así como también con respecto a algunos de los acontecimientos que se producen en nuestro cuerpo. Así, hay pensamientos obsesivos, cuyo origen tal vez sea incierto y de los que no nos podemos deshacer, pensamientos que nos atacan inesperadamente, cuando menos lo esperamos y pensamientos que corren de cualquier manera por nuestra cabeza.

Los pensamientos que nos acosan de estas formas no ocurren por alguna acción activa de nuestra parte. Sin duda, es tentador sugerir que no son en absoluto pensamientos que *nosotros pensamos*, sino más bien pensamientos que *encontramos* dentro de nosotros. Esto expresaría nuestra sensación de que si bien estos pensamientos son acontecimientos en la historia de nuestra mente, no participamos en forma activa para que ocurran. El verbo *pensar* puede connotar una actividad –como en "estoy pensando cuidadosamente en lo que dijiste"– y con respecto a este aspecto de su significado no podemos suponer que los pensamientos estén necesariamente acompañados por el proceso de pensar. No es incoherente, a pesar del aire de paradoja que pueda tener la frase, decir que un pensamiento que ocurre en mi mente puede ser algo que *pienso* o no. Esto puede entenderse casi de la misma manera que la afirmación, menos discordante, de que un acontecimiento que ocurre en mi cuerpo puede ser algo que *hago* o no.

No obstante, lo que quiero considerar en particular son las pasiones. El hecho de que la palabra *pasión* connote pasividad presenta una especie de obstáculo lingüístico para la cómoda aplicación a las pasiones de las distinciones de las que me he ocupado. Además, bien puede ser que este obstáculo sea más sustancial que una mera etimología. Sin embargo, creo que hay que hacer una distinción útil, por incómoda que sea su expresión, entre pasiones

con respecto a las que somos activos y aquellas con respecto a las que somos pasivos. Entre nuestras pasiones, igual que entre los movimientos de nuestro cuerpo, hay algunas cuyos principios motores están dentro de nosotros y otras cuyos principios motores nos son externos.

Al parecer, Terence Penelhum niega esto. Según él, cualquier persona que trata de representar sus deseos como externos a ella cae en una "forma de artimaña moral" que incluye una "falsedad literal y flagrante".[2] Este autor condena cualquier representación de esta clase por evasiva y falsa, porque "niega que algún deseo [...] forme parte de la historia que cada uno está viviendo, cuando forma parte de ella" (p. 671). Penelhum reconoce que cuando una persona tiene un deseo que preferiría no tener, o cuando es inducido a actuar por un deseo que no quiere que lo induzca a actuar, puede sentir que su deseo es, de alguna manera, ajeno a ella. En estos casos, se debe admitir que tal vez la persona diga que no se identifica con el deseo. Sin embargo, Penelhum sostiene que el deseo con el que una persona no se identifica es "tan parte de ella como el deseo con el que se identifica" (p. 672). Su argumento para esta afirmación es simple: cada deseo debe, después de todo, pertenecer a *alguien*, y un deseo con el que una persona no se identifica claramente no pertenece a nadie más (p. 674).

Tratar el tema de esta manera me parece algo precipitado. En mi opinión, no es tan inequívocamente obvio que el deseo humano deba ser el deseo de alguna persona. Por lo menos, creo que podemos decir que existe un sentido interesante, distinto del sentido en el que esto *es* bastante obvio, en el que no es obvio en absoluto. Supongamos que en un vehículo atestado en movimiento una persona es empujada contra otra, y que la segunda persona pregunta por simple curiosidad quién la empujó; entonces sería bastante sensato que la primera persona dijera que fue ella, y que el asunto quedara ahí, aunque el empujón no hubiera supuesto una actividad de su parte. Un movimiento del cuerpo humano, aun cuando

2 T. Penelhum, "The importance of self-identity", en *Journal of Philosophy*, LXVIII, 1971, p. 670. Los números entre paréntesis que aparecen en el resto del párrafo se refieren a las páginas de este ensayo.

es un simple suceso en la historia de la persona cuyo cuerpo se mueve, puede identificarse, para ciertos propósitos y en forma adecuada, como un movimiento de esa persona y de nadie más. Sin embargo, nos parece útil reservar un sentido en el que un movimiento de esta clase no es en absoluto estrictamente atribuible a la persona, sino sólo a su cuerpo. Reconocemos que en este sentido estricto, no hay *persona alguna* a la que pueda atribuírsele, ninguna persona de la que sea "tan parte de ella" como lo son sus acciones y sus actividades. Ahora bien, ¿por qué un deseo no puede ser, de un modo similar, un acontecimiento en la historia de la mente de una persona sin ser el deseo de esa persona? ¿Por qué, en este sentido, no podría ser que ciertos movimientos mentales, como ciertos movimientos corporales, no pertenecieran a nadie?

Pensamos que es correcto atribuirle a una persona, en el sentido estricto, sólo algunos de los acontecimientos de la historia de su cuerpo. Los otros –aquellos respecto de los que es pasiva– tienen sus principios motores fuera de ella, y no la identificamos con estos acontecimientos. De la misma manera, ciertos acontecimientos en la historia de la mente de una persona tienen sus principios motores fuera de ella. La persona es pasiva respecto de ellos y, del mismo modo, ellos no deben serle atribuidos. En otras palabras, una persona no debe identificarse con todo lo que sucede en su mente más de lo que debe identificarse con todo lo que sucede en su cuerpo. Por supuesto, cada movimiento del cuerpo de una persona es un acontecimiento de su historia; en este sentido, es su movimiento y de nadie más. En este mismo sentido, todos los acontecimientos de la historia de la mente de una persona también son suyos. Si esto es todo lo que se quiere decir, entonces es innegable que una pasión no puede ocurrir sin pertenecer a alguien, así como un movimiento de un cuerpo humano viviente no puede ocurrir sin ser el movimiento de alguien. Sin embargo, esto es sólo una verdad literal y flagrante, que oculta las distinciones que son tan valiosas en un caso como lo son en el otro.

Las críticas de Penelhum son demasiado severas. Hacen caso omiso del hecho de que en la relación de una persona con sus pasiones hay problemas análogos a los problemas, más conocidos, que conciernen a la relación de una persona con los movimientos

de su cuerpo. Insistir en forma inequívoca en que toda pasión debe ser atribuible a alguien es, por tanto, tan gratuito como sería insistir en que un movimiento espasmódico del cuerpo de una persona es producido por esa persona, a menos que haya otra persona de la que se pueda decir que es *ella* la que realiza el movimiento. De hecho, un sentido legítimo e interesante en el que una persona puede experimentar una pasión externa a ella, y que no es estrictamente atribuible ni a ella ni a ningún otro.

Reconocer esto no necesariamente nos impide concordar con Penelhum en que dicha pasión es parte de la historia que la persona está viviendo. Más aun, puede observarse que negarse a atribuirle a una persona ciertas pasiones que experimenta no nos compromete a considerar esas pasiones totalmente irrelevantes para alcanzar un juicio justo con respecto a lo que podemos esperar de esa persona. Una pasión no es menos genuina, y su embestida no es menos potente, por ser externa a la persona en cuya historia ocurre, de la misma manera que la ocurrencia o los efectos de un movimiento corporal no son menos palpables por el hecho de que el movimiento no sea realizado por la persona en cuyo cuerpo ocurre.

Sin duda, estaríamos brindando a las personas oportunidades para la evasión moral, como sugiere Penelhum, si admitiéramos que puede ser legítimo que alguien niegue ciertas pasiones suyas por ser externas. Porque, por supuesto, una persona puede estar actuando de mala fe cuando niega que una pasión que encuentra en su interior debe atribuírsele inequívocamente a ella. Pero solemos admitir una actitud parecida de evasión y falta de autenticidad, sin que lo consideremos un error, cuando aceptamos que se puedan negar ciertos movimientos corporales por ser externos. Después de todo, una persona puede intentar, en forma deshonesta y exitosa, escapar a un juicio desfavorable al que, de otra manera, sería sometido negando que realizó determinado movimiento de su cuerpo, y profesando que el principio motor del acontecimiento físico en cuestión en realidad fue externo a ella. Además, puede resultar tan descorazonadoramente difícil descubrir el autoengaño o la mentira cuando alguien finge que no realizó un movimiento de su cuerpo, como suele serlo descubrir la

falta de sinceridad de una persona cuando sostiene que no debe atribuírsele una pasión que experimenta. Por cierto, la ambición de facilitar nuestra tarea de jueces morales no es una razón mejor en un caso que en el otro para pasar por alto una distinción que corresponde a una significativa diferencia de hechos.

II

Ahora debemos considerar cuáles de las pasiones en la historia de una persona *son* externas a ella y examinar las condiciones de su externalidad. Es particularmente probable que una pasión sea externa cuando es artificialmente inducida por medios tales como la hipnosis o el uso de drogas. En casos como éstos, la pasión no suele surgir como respuesta a una experiencia percibida. En consecuencia, puede presentársele a la persona en cuya historia se las arregla para aparecer como discontinua respecto de la comprensión que tiene de su situación y respecto de la concepción que tiene de sí misma. Aun así, muchas veces parece que la persona, por instinto, esquiva estas discontinuidades mediante la racionalización: instantáneamente proporciona significado a la pasión, o de alguna manera interpreta que tiene un lugar natural en su experiencia. Entonces, a pesar de su origen, la pasión se acopla a un principio motor dentro de la persona, y la persona ya no es más un observador pasivo con respecto a ella de lo que habría sido si la pasión hubiera surgido en una respuesta más integral a sus percepciones.

Sin embargo, las pasiones externas no son sólo las que se suscitan mediante artimañas. Consideremos el siguiente episodio ejemplar:

> En el curso de una conversación animada aunque bastante amistosa, un hombre de pronto pierde el control. Aunque nada ha pasado que haga que su conducta sea fácilmente inteligible, comienza a arrojar platos, libros y emplea un lenguaje groseramente vulgar con su acompañante. Luego se le pasa el berrinche y dice: "No tengo idea de qué puede haberme provocado ese

extraño espasmo de emoción. Los sentimientos me invadieron no sé de dónde, y no pude evitarlo. No era yo mismo. Por favor, no me guardes rencor".

Estos argumentos para negar la responsabilidad pueden ser, por supuesto, mecanismos vilmente hipócritas para obtener una indulgencia no merecida. O pueden no ser más que expresiones enfáticas de arrepentimiento. Sin embargo, también es posible que sean genuinamente descriptivos. Lo que el hombre dice bien puede estar transmitiendo su sensación de que el aumento de la pasión representó, de alguna manera, una intrusión en su interior, que lo violó, que cuando fue poseído por la ira no era dueño de sí mismo. En afirmaciones como las del hombre del ejemplo y en la sensación de sí mismo que dichas afirmaciones expresan, encontramos muy vívidamente la experiencia de la externalidad.

¿Qué características de los episodios como éste son condiciones esenciales o marcas de externalidad, y cuáles carecen de importancia? La respuesta que viene más fácilmente a la mente es que las pasiones son externas a nosotros cuando preferimos no tenerlas, o cuando preferimos no ser inducidos por ellas; y que son internas cuando, en el momento en que ocurren, las recibimos con gusto o las aceptamos con indiferencia. A causa de esto, una pasión es inequívocamente nuestra cuando es lo que queremos sentir, o estamos dispuestos a sentir, mientras que una pasión cuya ocurrencia en nuestro interior desaprobamos no es estrictamente nuestra.

En ocasiones, cuando desaprobamos el curso que tomaron nuestras pasiones, decimos que no representan lo que sentimos *realmente*. No lo hacemos con la intención de negar que las pasiones en cuestión estén ocurriendo, sino con la de indicar que consideramos que de algún modo son incoherentes con la concepción de nosotros mismos que preferimos, la que suponemos que captura lo que somos con más autenticidad que una mera descripción no destilada. Con frecuencia, las personas tienden –por lo menos hasta que llegan a cierta edad– a interpretar lo que realmente son como lo que les gustaría ser. Consideran que sus *verdaderas* pasiones son aquellas por las que les gustaría sentirse motivadas, o con

las que preferirían que las identificaran quienes las conocen, aunque, de hecho, estas pasiones puedan ser más débiles e influir menos en su comportamiento que otras.

Esta ecuación de lo real con lo ideal desempeña un papel en la forma en que ciertas personas piensan de sí mismas. Sin embargo, la distinción entre pasiones internas y externas no es la misma que la distinción entre lo que es y lo que no es *real* en el sentido de ajustarse a la imagen ideal que una persona tiene de sí misma. Sin duda, es posible que una persona reconozca que cierta pasión se le puede atribuir inequívocamente a ella, aun cuando lamente este hecho y aun cuando hubiera deseado que la pasión no se produjera en ella o no la indujera en absoluto. Tal vez luego de una larga lucha y desilusión consigo misma, una persona pueda resignarse a ser alguien que ella misma no termina de aprobar. Ya no supone que es capaz de lograr que el curso de sus pasiones esté en armonía con el concepto ideal que tiene de sí misma y, como consecuencia, deja de retener la aceptación de sus pasiones tal como son.

Tampoco es esencial para la externalidad de una pasión que sea de irresistible intensidad. En general, solemos negar una pasión sólo cuando resulta que habríamos preferido no estar sometidos a ella, o cuando interfiere de alguna forma importante con la soberanía de otras pasiones que consideramos más genuinamente nuestras. Las consideraciones de este tipo explican nuestro interés en *llamar la atención hacia* la externalidad de ciertas pasiones, pero no son condiciones de la externalidad misma. No hay nada en la noción de externalidad que suponga un carácter irresistible o, para el caso, cualquier nivel particular de intensidad relativa. Así, es muy comprensible que una persona encuentre dentro de sí un deseo con el que no se identifica y, sin embargo, que no tenga ninguna dificultad para impedir que ese deseo la induzca a actuar o usurpe su voluntad. En mi ejemplo, la persona lamenta su incapacidad para controlar su ira. Esto explica su anhelo de disociarse públicamente de esa pasión y de los deseos a los que la condujo. No obstante, ni su incapacidad ni su arrepentimiento explican el hecho –si es que fue un hecho– de que la ira y los airados deseos no hayan sido estrictamente suyos.

III

He sostenido que la pregunta de si una pasión es interna o externa a una persona no es simplemente una cuestión de la actitud de la persona hacia la pasión. Aun así, me resisto a sugerir que las actitudes de una persona hacia sus pasiones pueden desecharse por ser totalmente irrelevantes para esta cuestión. Supongamos que alguien niega una pasión que ocurre dentro de ella en cierta ocasión, pero que nosotros sabemos que suele experimentar pasiones de esta clase en ocasiones como ésa y que se muestra muy dispuesta y satisfecha de que así sea. Es probable que, en ese caso, no tratemos la negación de su responsabilidad con mucha seriedad, y parece una respuesta razonable. Lo que no queda claro es cuál es el fundamento que nos permite ser escépticos con razón.

Quizá pensamos que el hecho de si una persona debe identificarse o no con cierta pasión depende de ella, por lo menos en ocasiones. En otras palabras, quizá tenemos la idea de que una persona puede actuar de algún modo para hacer que una pasión sea completamente suya, y que a veces no hay nada que le impida hacerlo si decide hacerlo. Entonces nos parecería difícil entender, cuando una persona experimenta cierta pasión con satisfacción y una regularidad predecible, que no haya hecho lo que fuera necesario para identificarse con ella. Eso explicaría nuestro escepticismo.

El hecho de que una persona desapruebe una pasión que ocurre dentro de ella no es, en sí, equivalente al hecho de que la pasión sea externa. Pues es posible, como ya he observado, que alguien se resigne a lo que considera sus defectos. Una persona puede reconocer que las pasiones que desaprueba son innegable e inequívocamente suyas; y entonces puede dejar de sentir, si alguna vez lo sintió, que estas pasiones le son de alguna manera ajenas o que la invaden. El hecho de que una persona desapruebe una pasión no es, por tanto, una condición suficiente de la externalidad de esa pasión. Por otro lado, puede ser que la desaprobación sea una condición necesaria de la externalidad. De hecho, es difícil pensar en un ejemplo convincente en el que una per-

sona para quien una pasión es externa apruebe, de todas maneras, que una pasión ocurra dentro de ella. Y la dificultad de encontrar un ejemplo de esta clase también tiende a respaldar la conjetura de que la aprobación de una persona de una pasión que ocurre en su historia es una condición suficiente de que la pasión sea interna a ella.

Cualquiera sea la verdad acerca de estas relaciones, es importante notar que existe un error básico en pensar que los conceptos de internalidad y externalidad deben ser explicados *simplemente* en términos de las actitudes de una persona. Fundamentalmente, es erróneo sugerir que la externalidad de una pasión es consecuencia de la desaprobación de una persona, o que su internalidad es consecuencia de su aprobación. La dificultad que plantea este acercamiento al problema de entender la internalidad y la externalidad es que no tiene en cuenta el hecho de que las actitudes hacia las pasiones son tan susceptibles de la externalidad como las pasiones mismas. Supongamos que una persona tiene sentimientos encontrados con respecto a una de sus pasiones: es consciente de sentir una inclinación a aprobar la pasión y también de sentir una inclinación a desaprobarla. Supongamos que resuelve su conflicto adoptando una decidida actitud de desaprobación respecto de la pasión. Sin embargo, puede encontrar que su inclinación a aprobar la pasión persiste, aunque ahora le es externa y no puede atribuírsele como propia.

Se puede interpretar que el hecho de que una persona tenga cierta actitud hacia una pasión determina ya sea la internalidad, ya la externalidad de la pasión, por cierto, sólo si la actitud en cuestión es genuinamente atribuible a la persona. La actitud en virtud de la cual una pasión es interna o externa no puede estar entre las que una persona simplemente encuentra en su interior; debe ser una actitud con la que la persona se identifica. Sin embargo, dado que la cuestión de la atribución surge no sólo con respecto a las pasiones de una persona, sino también con respecto a sus actitudes hacia sus pasiones, se generará un regreso al infinito cuando se intenta explicar la internalidad o la externalidad en términos de actitudes. Ello se debe a que la actitud que se invoca para explicar el estatuto de una pasión tendrá que ser

interna; su internalidad deberá explicarse invocando una actitud de orden superior, es decir, una actitud hacia una actitud; la internalidad de esta actitud de orden superior deberá explicarse en términos de una actitud de un orden aun superior, y así sucesivamente. Esto impide que los conceptos de internalidad y externalidad sean explicados simplemente mediante la noción de la jerarquía de las actitudes.

IV

No puedo proporcionar una descripción satisfactoria de lo que significa caracterizar una pasión como interna o como externa. Sin embargo, quizá pueda aclarar estas nociones un poco más analizando dos tipos de conflicto del deseo. Supongamos que una persona quiere ir a un concierto y que también quiere ver una película. Imaginemos también que, dadas las circunstancias, estos deseos entran en conflicto, de modo que la persona debe decidir si prefiere pasar la tarde de un modo o de otro. Ahora bien, es probable que tanto el deseo de ver la película como el deseo de ir al concierto sean igualmente internos a la persona. Se trata de una situación en la que la persona simplemente quiere hacer las dos cosas y no puede. El modo de resolver el problema que presenta este conflicto sería decidir cuál de las dos cosas prefiere hacer.

Los conflictos de este tipo sólo requieren que los deseos en cuestión sean jerarquizados. Su rasgo esencial es que los deseos conflictivos pertenecen a la misma jerarquía, aunque mientras el conflicto no se resuelva, no se establecen sus posiciones relativas en esa jerarquía. La pertenencia a la misma jerarquía puede explicarse de la siguiente manera: supongamos que la persona decide que en realidad prefiere ver la película, pero resulta que no consigue entradas. Entonces sería muy natural que se inclinara por su segunda opción y fuera al concierto. Después de todo, también quiere ir allí. Su decisión original de no ir sólo significa que satisfacer su deseo de ir al concierto tenía menor jerarquía que satisfacer su deseo de ver la película.

Cuando alguien sufre un conflicto del segundo tipo, su problema no es ubicar un deseo que siente en una posición inferior a la de otro en una escala única. Es rechazar uno de sus deseos en forma absoluta. Los conflictos entre deseos conscientes e inconscientes son característicamente de este tipo, pero ambos deseos conflictivos bien pueden permanecer conscientes. Supongamos que una persona quiere felicitar a un conocido por algún logro reciente, pero que también nota en su interior un deseo celosamente rencoroso de herirlo. Puede ser que, tal como resultan las cosas, no encuentre la oportunidad de hacer el comentario amistoso que (imaginamos) había decidido hacer. Esto no la llevaría naturalmente a ver si puede lograr la satisfacción de su otro deseo. Cuando alguien se frustra en su deseo de ver una película, naturalmente se inclina hacia su segunda opción y va al concierto. En cambio, en este ejemplo, la alternativa de herir a su conocido no es la segunda de las opciones de la persona, cuya primera opción era la de felicitarlo. En este sentido, los deseos originados en la amistad y en los celos, a diferencia de los que se refieren al concierto y a la película, no pertenecen a la misma jerarquía.

No hay razón para suponer que el deseo de la persona de herir a su conocido se produzca en primera instancia como un deseo externo, y que su deseo de felicitarlo se produzca originalmente como uno interno. Sin duda, puede suceder de esa manera. Si así fuera, entonces los dos deseos se presentarán desde el principio como pertenecientes a distintas jerarquías; y la persona no tendrá conflicto que resolver, aunque tal vez deba luchar contra la influencia del deseo externo. Por otro lado, puede suceder que, en una etapa temprana de la confrontación de sus sentimientos, la persona encuentre que un deseo no es para ella menos interno que el otro. En ese caso, la persona misma está en conflicto, no sólo los deseos que están ocurriendo en su historia mental. Tiene el problema de no saber qué quiere hacer, en un sentido conocido que es bastante compatible con el hecho de que sabe que quiere tanto herir a su conocido como felicitarlo, y el conflicto que manifiesta su incertidumbre requiere una definición.

La persona no puede resolver este conflicto decidiendo que su primera opción es satisfacer uno de los deseos en conflicto y que su

segunda opción es satisfacer el otro deseo. En otras palabras, este conflicto no se resuelve *jerarquizando* los deseos en conflicto sino *rechazando* uno de ellos. Al rechazar el deseo de herir a su conocido, suponiendo que esto sea lo que hace, la persona se aleja de ese deseo. Coloca el deseo rechazado fuera del alcance de sus preferencias, de modo que deja de ser una opción que podrá ser satisfecha. Aunque es posible que el deseo rechazado siga estando presente en su historia mental, de esta manera la persona hace que su ocurrencia sea externa. Entonces, el deseo ya no debe atribuírsele estrictamente a ella, aunque pueda persistir o ser recurrente como un elemento de su experiencia.

En un conflicto del primer tipo, la decisión de la persona de ver la película no significa que no quiera ir al concierto. Significa que no quiere ir al concierto *tanto* como ir a ver la película. En el segundo, el rechazo de la persona del deseo de herir a su conocido no significa simplemente que quiera felicitarlo *más* de lo que desea herirlo. Significa que *no* quiere herirlo, aunque experimente el deseo de herir. Sería un poco engañoso insistir en que, dado que el deseo de felicitar a su conocido prevalece sobre el deseo de herirlo, se da simplemente el caso que el primero es más fuerte que el segundo. Esto sugiere equivocadamente que ambos deseos ocupan posiciones en la misma jerarquía y que sólo difieren en fuerza. Creo que hay una forma mejor de describir el hecho de que el deseo de felicitar prevalezca sobre el deseo de herir: decir que *la persona*, que quiere felicitar a un conocido, es más fuerte que el deseo de herirlo que encuentra en su interior.

Al decidir que lo que quiere, después de todo, es felicitar a su conocido, y que su deseo de herir al hombre finalmente debe excluirse del orden de los deseos susceptibles de ser satisfechos, la persona hace que el segundo deseo sea externo a ella y se identifica con el primero. Aquí parece que la relación de la persona con sus pasiones queda establecida por el hecho de tomar un tipo particular de decisión. Tal vez, una decisión de esta clase, aun cuando no sea tan visible como en este ejemplo, está detrás de cada instancia de la definición de la internalidad o la externalidad de las pasiones. O quizá debamos tratar de entender los fenómenos en cuestión haciendo referencia a algo más general, de lo que las de-

cisiones son sólo casos especiales. De cualquier modo, la naturaleza de la decisión es muy poco clara.[3]

V

El problema de la atribución o de la externalidad con respecto a los movimientos corporales es explicar qué diferencia un movimiento que hace una persona de un mero suceso en la historia de su cuerpo. El problema correspondiente a los fenómenos psíquicos concierne a la naturaleza de la identificación: ¿qué significa para una persona ser identificada con una pasión más que con otra (o lo que sea), cuando ambas son suyas en el burdo sentido literal? Cada uno de estos problemas es, en su propio ámbito, fundamental y frustrantemente recalcitrante. He hecho poco más que sugerir que, en cierto sentido, son análogos. Queda por ver en qué modos importantes pueden diferir y si es probable que sean susceptibles de soluciones análogas.

3 A diferencia de los deseos o las actitudes, las decisiones no parecen ser susceptibles de la internalidad y de la externalidad. Mencionarlas aquí parecería evitar, por consiguiente, la dificultad considerada en la sección precedente.

6
El problema de la acción

I

El problema de la acción reside en explicar el contraste entre lo que hace un agente y lo que meramente le sucede, o entre los movimientos corporales que realiza y aquellos que se producen sin que él los provoque. De acuerdo con las teorías causales sobre la naturaleza de la acción –en la actualidad, el enfoque más generalizado en la comprensión de este contraste–, la diferencia esencial entre los acontecimientos de ambos tipos se encuentra en sus historias causales previas: un movimiento corporal es una acción si y sólo si es el resultado de antecedentes de cierto tipo. Las diversas versiones del enfoque causal pueden proporcionar descripciones diferentes de los tipos de acontecimientos o de estados que deben figurar causalmente en la producción de acciones. El principio característico que comparten es que considerar cómo se produjo un acontecimiento es tanto necesario como suficiente para determinar si ese acontecimiento es una acción.

A pesar de su popularidad, creo que el enfoque causal es intrínsecamente poco razonable, y que no puede brindar un análisis satisfactorio de la naturaleza de la acción. No intento sugerir que las acciones carezcan de causas; es tan probable que tengan causas, supongo, como otros acontecimientos. Sostengo, en cambio, que no es parte de la naturaleza de una acción tener una historia causal previa de algún tipo particular. En mi opinión, del hecho de que un acontecimiento sea una acción no se desprende siquiera que

tenga una causa o causas ni, mucho menos, antecedentes causales de algún tipo específico.

Cuando afirman que la diferencia esencial entre las acciones y los meros sucesos reside en sus historias causales previas, las teorías causales presuponen que en esencia las acciones y los meros sucesos no difieren. Estas teorías sostienen que las secuencias causales que producen acciones necesariamente son de un tipo diferente de aquellas que producen meros sucesos, pero que los efectos producidos por las secuencias de ambos tipos son, desde un punto de vista intrínseco, indistinguibles. Por tanto, están obligadas a suponer que una persona que sabe que está en medio de la realización de una acción no puede haber obtenido ese conocimiento de la conciencia de lo que está sucediendo en ese momento, sino que lo debe de haber obtenido, en cambio, de su comprensión acerca de cómo lo que está sucediendo fue causado por ciertas condiciones previas. Para el enfoque causal, es esencial considerar que las acciones y los meros sucesos no se diferencian por nada que existe o que sucede en el momento en que esos acontecimientos ocurren, sino por algo que les es extrínseco: una diferencia, en un momento previo, en otra serie de acontecimientos.

Esto es lo que hace que las teorías causales sean poco razonables. Alejan la atención exclusivamente de los acontecimientos cuya naturaleza se analiza y de los momentos en que ocurren. El resultado es que queda más allá de su alcance estipular que la persona debe relacionarse de algún modo específico con los movimientos de su cuerpo *durante* el período en que se presume que está llevando a cabo una acción. Las únicas condiciones que, según estas teorías, son distintivamente constitutivas de la acción pueden dejar de imperar, debido a todo lo que exigen las explicaciones causales, precisamente cuando el agente comienza a actuar. Una vez que han ocurrido los antecedentes causales de la realización de la acción, no exigen nada de un agente excepto que su cuerpo se mueva como efecto de esos antecedentes.

No es sorprendente que esas teorías suelan toparse con contraejemplos de un tipo bien conocido. Por ejemplo: un hombre que se encuentra en una fiesta tiene la intención de derramar el contenido de su copa como señal para sus cómplices de que de-

ben iniciar un robo y cree, en virtud de sus arreglos previos, que derramando lo que tiene en su copa logrará su objetivo; pero todo esto hace que el hombre se sienta muy ansioso, su ansiedad le hace temblar la mano y, como consecuencia, derrama el contenido de la copa. Independientemente del tipo de antecedentes causales que se consideren necesarios y suficientes para que ocurra una acción, resulta fácil demostrar que los antecedentes causales de este tipo pueden tener como efecto un acontecimiento que, de manera manifiesta, no sea una acción, sino un simple movimiento corporal. En el ejemplo dado, entre las causas que hacen que la copa se derrame se encuentran un deseo y una creencia, que racionalizan el hecho de que el hombre derrame el contenido de su copa; pero la forma en que se derrama, tal como sucede, no constituye una acción. Ese ejemplo, en particular, le crea problemas a una teoría causal en la que se interpreta que las acciones son, en esencia, movimientos cuyas causas son deseos y creencias mediante los cuales se racionalizan. Es fácil imaginar contraejemplos de esa índole para crearles problemas similares a otras variantes del enfoque causal.

No analizaré las diversas maniobras mediante las cuales los teóricos causales han intentado hacer frente a estos contraejemplos.[1] En mi opinión, las teorías causales son, en forma inevitable, vulnerables a dichos contraejemplos, porque localizan las características distintivas de la acción sólo en situaciones que pueden haber pasado para cuando se supone que la acción debe tener lugar. Esto les imposibilita dar algún tipo de explicación acerca de la característica diferenciadora más destacada de la acción: durante el

1 Para conocer algunos análisis del problema según los adherentes al enfoque causal, pueden consultarse Alvin Goldman, *A theory of human action*, Princeton, 1970, pp. 61-63; Donald Davidson, "Freedom to act", en T. Honderich (comp.), *Essays on freedom of action*, Londres, 1973, pp. 153-154; Richard Foley, "Deliberate action", en *The Philosophical Review*, vol. 86, 1977, pp. 58-69. Es evidente que Goldman y Davidson creen que el problema de evitar los contraejemplos es empírico y que corresponde que sea trasladado a los científicos. La "solución" de Foley renuncia a la obligación de brindar un análisis adecuado de otra manera: especifica las condiciones para actuar y, cuando reconoce que pueden ser satisfechas con espasmos y tics, simplemente declara que dichos movimientos son, no obstante, acciones si satisfacen sus condiciones.

tiempo en que una persona está realizando una acción, necesariamente está en contacto con los movimientos de su cuerpo en cierta manera, mientras que necesariamente no está en contacto con ellos, de esa manera, cuando los movimientos de su cuerpo ocurren sin que ella los provoque. Una teoría que se limita a describir las causas previas a la ocurrencia de las acciones y de los meros movimientos corporales de ningún modo puede incluir un análisis de estas dos maneras en las que una persona puede relacionarse con los movimientos de su cuerpo. Inevitablemente tiene que dejar abierta la posibilidad de que una persona, cualquiera que sea su participación en los acontecimientos de los que surgen sus acciones, pierda toda conexión con los movimientos de su cuerpo en el momento en que comienza la acción.

II

A fin de elaborar una forma más prometedora de pensar acerca de la acción, consideremos la noción de que las acciones y los meros sucesos son indistinguibles. Esta noción es un elemento importante en la motivación de las teorías causales. Si se pensara que las acciones y los meros sucesos difieren intrínsecamente, sería obvio que la manera de explicar cómo difieren sería identificando esta diferencia intrínseca entre ellos. Debido a que los teóricos causales piensan que no hay otra manera de diferenciar las acciones de los meros sucesos, buscan una diferencia diferenciadora entre los acontecimientos que los preceden.

David Pears, quien cree que los deseos desempeñan un papel esencial en la producción de las acciones, lo expresa de la siguiente manera:

> Simplemente, no poseemos la capacidad general de distinguir entre aquellos movimientos corporales que son acciones y aquellos que son meros movimientos corporales sin usar como criterio la presencia o la ausencia del deseo pertinente. [...] Es cierto que hay diversas características intrínsecas de los movi-

> mientos corporales que brindan algún indicio para su clasificación. Por ejemplo, es probable que un movimiento muy complicado haya sido producido por un deseo. Pero [...] la simplicidad de un movimiento ni siquiera hace probable que no haya sido producido por un deseo.

Debido a que no podemos hallar ninguna característica inherente de la acción que nos permita distinguirla con seguridad de los meros movimientos corporales, debemos, por tanto, según la opinión de Pears, "clasificar algunos movimientos corporales como acciones solamente en virtud de sus orígenes".[2]

Pears observa correctamente que los movimientos del cuerpo de una persona no revelan en forma definitiva si está llevando a cabo una acción: los mismos movimientos pueden tener lugar cuando se está llevando a cabo una acción o cuando está ocurriendo un mero suceso. No obstante, de esto no se desprende que la única manera de descubrir si una persona está actuando o no es considerar qué estaba sucediendo *antes* de que comenzaran sus movimientos: es decir, considerando las causas que los originaron. De hecho, la situación que tiene lugar *mientras* ocurren los movimientos es mucho más pertinente. Lo que no es meramente pertinente, sino decisivo, es, sin duda, considerar si los movimientos *son guiados por la persona* cuando ocurren o no. Esto es lo que determina si la persona está llevando a cabo una acción. Más aun, la cuestión de si los movimientos ocurren o no guiados por una persona no se relaciona con sus antecedentes. Los acontecimientos son causados por situaciones precedentes, pero un acontecimiento no puede ser guiado en el curso de su ocurrencia a una distancia temporal.

Vale la pena notar que Pears está equivocado cuando sostiene que los movimientos muy complicados, aunque puedan ser meros sucesos, probablemente deban clasificarse como acciones. Sin duda, los movimientos complicados de las manos y los dedos de un pianista sugieren, en forma convincente, que no son meros sucesos. A

2 D. Pears, "Two problems about reasons for actions", en R. Binkley, R. Bronaugh, A. Marras (comps.), *Agent, action and reason*, Oxford, 1971, pp. 136-137, 139.

veces, sin embargo, la complejidad puede sugerir, en forma igualmente convincente, la posibilidad de un mero movimiento corporal. Las sacudidas que sufre el cuerpo de una persona durante un ataque de epilepsia, por ejemplo, son movimientos muy complicados. Pero su complejidad es de un tipo que nos hace parecer improbable que la persona esté llevando a cabo una acción.

¿En qué circunstancias la complejidad de los movimientos sugiere la acción, y en cuáles, su ausencia? Esto depende, por así decirlo, de si los movimientos en cuestión se unen para crear un patrón que nos parezca significativo. Cuando esto sucede, como en el caso del pianista, nos resulta difícil imaginar que los movimientos podrían haber ocurrido, en la complicada forma que requiere el patrón significativo que han creado, si el pianista no hubiera guiado sus manos y sus dedos cuando éstos se movían. En el caso del epiléptico, en cambio, nos parece improbable que una persona pudiera haber creado un patrón tan incoherentemente complicado si hubiese estado guiando su cuerpo para que realizara esos movimientos. Los movimientos simples de una persona, como observa Pears, por lo general no sugieren ni una acción ni un mero suceso. Ello se debe a que, normalmente, sus patrones no nos parecen, en sí mismos, ni significativos ni incoherentes. No nos brindan, en apariencia, ningún indicio de si están siendo guiados o no por la persona en el momento en que ocurren.

La complejidad del movimiento corporal sugiere la acción sólo cuando nos conduce a pensar que el cuerpo, durante el curso de su movimiento, es guiado por el agente. La realización de una acción es, por tanto, un acontecimiento complejo, que está compuesto por un movimiento corporal y por cualquier situación o actividad resultante de la guía que el agente le proporciona a la acción. Dado un movimiento corporal que ocurre guiado por una persona, la persona está llevando a cabo una acción independientemente de las características de su historia causal previa que expliquen el hecho de que esto esté ocurriendo. La persona está llevando a cabo una acción incluso si la ocurrencia de ésta se debe a la casualidad. Y no está llevando a cabo una acción si los movimientos no son guiados por ella a medida que ocurren, incluso si la persona misma hubiera proporcionado las causas precedentes –en la forma de creencias,

deseos, intenciones, decisiones, voliciones o lo que fuere–, de las cuales provino el movimiento.

III

Cuando actuamos, nuestros movimientos tienen un propósito. Ésta es sólo otra forma de decir que su curso está guiado. Muchos casos de movimientos con un propósito no son, por supuesto, instancias de acción. La dilatación de las pupilas de los ojos de una persona cuando disminuye la luz, por ejemplo, es un movimiento con un propósito; hay mecanismos que guían su curso. Sin embargo, la ocurrencia de este movimiento no marca la realización de una acción por parte de la persona; sus pupilas se dilatan, pero ella no las dilata. Esto se debe a que el curso del movimiento no está guiado *por la persona*. En este caso, sólo se puede atribuir la guía a la operación de algún mecanismo con el que no se puede identificar a la persona.

Empleemos el término *intencional* para referirnos a los casos de movimientos con propósito en los que el agente proporciona la guía. Podemos decir, entonces, que la acción es movimiento intencional. La noción de movimiento intencional no debe confundirse con la de acción intencional. El término *acción intencional* puede emplearse o, más bien, emplearse mal, sólo para expresar que una acción es necesariamente un movimiento cuyo curso es guiado por un agente. Cuando se lo emplea de esta manera, el término es pleonástico. Una forma apropiada de usarlo sería para hacer referencia a las acciones que son llevadas a cabo en forma más o menos deliberada o consciente; es decir, a las acciones que el agente tiene la intención de llevar a cabo. En este sentido, las acciones no son necesariamente intencionales.

Cuando una persona tiene la intención de llevar a cabo una acción, tiene la intención de que ocurran ciertos movimientos intencionales de su cuerpo. Cuando estos movimientos ocurren, la persona está llevando a cabo una acción intencional. Se podría decir, entonces, que está guiando los movimientos de su cuerpo

de cierta manera (por tanto, está actuando), y que, al hacerlo, la persona está siendo guiada por su intención de hacerlo y la está cumpliendo (por tanto, está actuando en forma intencional). Al parecer no hay nada en la noción de un movimiento intencional que implique que su ocurrencia dependa de la intención del agente, ya sea mediante una reflexión previa o un asentimiento consciente. Si esto es correcto, las acciones (es decir, los movimientos intencionales) pueden realizarse tanto de manera intencional como no intencional.

Debido a que la acción es movimiento intencional o un comportamiento cuyo curso es guiado por un agente, una explicación de la naturaleza de la acción debe abordar dos problemas diferentes. Uno es explicar la noción de comportamiento guiado. El otro es especificar cuándo la guía del comportamiento puede atribuirse a un agente y no, simplemente, como sucede cuando las pupilas de una persona se dilatan debido a que disminuye la luz, a algún proceso local que se produce dentro del cuerpo del agente. El primer problema se refiere a la condiciones en las cuales el comportamiento tiene un propósito, mientras que el segundo se refiere a las condiciones en las cuales el comportamiento con un propósito es intencional.

El conductor de un automóvil guía el movimiento de su vehículo mediante la acción: gira el volante, aprieta el acelerador, frena, etc. Nuestra guía de los movimientos que realizamos, cuando actuamos, no requiere, de manera similar, que realicemos diversas acciones. No tenemos el control de nuestros cuerpos de la manera en que un conductor tiene el control de su automóvil. De otro modo, no se podría concebir la acción –so pena de generar un regreso al infinito– como la ocurrencia de movimientos que son guiados por un agente. El hecho de que nuestros movimientos, cuando actuamos, tengan un propósito no es el efecto de algo que hacemos. Es una característica de la operación, en ese momento, de los sistemas que somos.

La conducta tiene un propósito cuando su curso está sujeto a las modificaciones que compensan los efectos de fuerzas que, de otra forma, interferirían en el curso de la conducta, y cuando la ocurrencia de estas modificaciones no puede explicarse mediante lo

que explica la situación que las provoca. En ese caso, la conducta está guiada por un mecanismo causal independiente, cuya disposición para originar modificaciones compensatorias tiende a asegurar que se cumpla la conducta.[3] Desde luego, la actividad de un mecanismo tal no suele estar guiada por nosotros. En cambio, cuando estamos llevando a cabo una acción, *es* la guía de nuestra conducta. La sensación que tenemos de nuestra propia agencia cuando actuamos no es nada más que lo que sentimos cuando, de alguna manera, estamos en contacto con la operación de mecanismos de este tipo, mediante los cuales nuestros movimientos son guiados y su curso, garantizado.

Explicar la conducta con un propósito en términos de mecanismos causales no equivale a proponer una teoría causal de la acción. En primer lugar, la actividad pertinente de estos mecanismos no es previa, sino concurrente con los movimientos que guían. Sin embargo, de todas formas, no es esencial para el propósito de un movimiento que sea afectado causalmente por el mecanismo bajo cuya guía se lleva a cabo el movimiento. Un conductor cuyo automóvil se desliza colina abajo sólo debido a las fuerzas de la gravedad puede sentirse completamente satisfecho con su velocidad y su dirección y, por tanto, puede que no intervenga en absoluto para modificar su movimiento. Esto no demostraría que el movimiento del automóvil no ocurrió bajo su guía. Lo que importa es que él estaba preparado para intervenir si hubiese sido necesario, y que estaba en la posición de hacerlo con mayor o menor efectividad. De manera similar, es posible que los mecanismos causales que están listos para afectar el curso de un movimiento corporal nunca tengan ocasión de hacerlo, ya que puede no producirse ninguna retroalimentación negativa del tipo que desencadenaría su actividad

3 Ernest Nagel proporciona un análisis útil sobre esta forma de comprender la conducta con un propósito: "Goal-directed processes in biology", en *The Journal of Philosophy*, vol. 74, 1977, pp. 271 y ss. Los detalles de los mecanismos en virtud de los cuales algún elemento de la conducta tiene un propósito pueden ser descubiertos, por cierto, sólo mediante la investigación empírica. Sin embargo, especificar las condiciones que un mecanismo como ése debe satisfacer es un problema filosófico que pertenece al análisis de la noción de conducta con propósito.

compensatoria. La conducta tiene un propósito, no porque sea el resultado de causas de cierto tipo, sino porque sería afectada por ciertas causas si la consecución de su curso corriera peligro.

IV

Dado que el hecho de que ciertas causas originan una acción se diferencia de las consideraciones en virtud de las cuales es una acción, en principio no hay razones por las cuales una persona no pueda ser llevada, de maneras diversas, a realizar la misma acción. Esto resulta importante en el análisis de la libertad. En general, se acepta que una persona actúa libremente sólo si podría haber actuado de otra manera. Sin embargo, algunos casos que implican cierto tipo de sobredeterminación proporcionan contraejemplos evidentes de este principio: el *principio de las posibilidades alternativas*. En estos casos, una persona lleva a cabo una acción únicamente por sus propias razones, lo que nos lleva a considerar que la ha realizado en forma libre; no obstante, de otra manera, habría sido llevada a realizarla por fuerzas ajenas a su voluntad, por lo que, en realidad, no puede evitar actuar como lo hace.[4]

Así, supongamos que un hombre consume heroína porque goza de sus efectos y los considera beneficiosos. Pero también supongamos que, sin saberlo, es adicto a la droga y, por tanto, será impulsado a consumirla en cualquier caso, incluso si sus propias creencias y actitudes no lo llevan a hacerlo. Entonces, parece que consume la droga libremente, que no podría haber hecho otra cosa que consumirla, y que, por tanto, el principio de las posibilidades alternativas es falso.

Donald Davidson argumenta lo contrario: que si bien una persona hace intencionalmente lo que hace por razones propias, no

4 Véanse mis ensayos "Posibilidades alternativas y responsabilidad moral" y "La libertad de la voluntad y el concepto de persona", ambos incluidos en este volumen.

hace intencionalmente lo que fuerzas ajenas lo llevan a hacer. Aunque los movimientos de su cuerpo puedan ser los mismos en ambos casos, Davidson sostiene que la persona no está llevando a cabo una acción cuando los movimientos ocurren independientemente de las actitudes y las creencias pertinentes. Alguien que actuó libremente podría haber hecho lo mismo incluso si no hubiese sido motivado por su cuenta a hacerlo, pero sólo en el sentido de que su cuerpo podría haber hecho los mismos movimientos: "No habría actuado intencionalmente si las condiciones actitudinales no hubiesen estado presentes". Entonces, incluso en los casos *sobredeterminados*, algo depende del agente: "No [...] lo que hace (cuando se lo describe de un modo que deja abierta la cuestión de si fue intencional), sino si lo hace intencionalmente".[5]

Aquí, el asunto no es, tal como sugiere Davidson en un momento, si la *acción* de una persona puede ser intencional cuando fuerzas ajenas, más que sus propias actitudes, explican lo que hace. Se trata de si su *conducta* puede ser intencional en esas circunstancias. Ahora bien, la conducta del adicto que no tiene conciencia de su adicción es claramente tan intencional cuando es llevado a consumir la droga por la fuerza compulsiva de su adicción como cuando la consume por libre elección. Sus movimientos no son meros sucesos cuando consume la droga, porque no puede evitarlo. Entonces, está llevando a cabo la misma acción que habría realizado si hubiera consumido la droga libremente y con la ilusión de que podría haber hecho otra cosa.

Este ejemplo no está destinado a mostrar que Davidson está equivocado cuando insiste en que no puede haber una acción sin intencionalidad o cuando no se cumplen las condiciones actitudinales pertinentes. Después de todo, aun cuando el adicto sea impulsado a hacer lo que hace, presumiblemente su conducta se ve afectada tanto por su anhelo de consumir la droga como por su creencia de que el procedimiento que sigue cuando la consume lo aliviará. Sus movimientos, cuando se clava la aguja de la jeringa en el brazo y presiona el émbolo, son, por cierto, intencionales. Sin embargo, el problema pertinente no es si una acción puede ocurrir

5 D. Davidson, *op. cit.*, pp. 149-150.

independientemente de las condiciones actitudinales. Se trata de si es posible que una acción sea causada sólo por fuerzas ajenas.

Esto parecerá imposible sólo si se piensa que entre las causas de una acción debe haber condiciones actitudinales. Sin embargo, para una acción no es esencial que tenga una historia causal previa de ningún tipo específico. Entonces, incluso si no puede haber acción en ausencia de ciertas condiciones actitudinales, estas condiciones no son esenciales en tanto causas previas. El ejemplo se relaciona con el problema en cuestión, ilustrando cómo una acción (que incluya, por supuesto, todos los constituyentes actitudinales necesarios) puede no tener causas, salvo las no actitudinales o ajenas. En consecuencia, confirma la falsedad del principio de las posibilidades alternativas, demostrando que una persona puede ser inducida, sólo por fuerzas ajenas, a llevar a cabo una acción que también podría llevar a cabo por sí sola.

Dicho sea de paso, el ejemplo también sugiere que las condiciones actitudinales de la acción de una persona pueden, en sí mismas, serle ajenas. No hay razones para suponer que un adicto que sucumbe a su anhelo contra su voluntad finalmente adopte como propio el deseo al que ha intentado resistirse. Es posible que al final sólo se someta a él con resignación, como un hombre que sabe que ha sido vencido y que, por tanto, acepta con desesperación las consecuencias que la derrota le acarreará y no como alguien que decide unirse a las fuerzas a las que antes se había opuesto. Asimismo, hay creencias obsesivas y falsas, por ejemplo: "Romper un espejo trae siete años de mala suerte", que una persona puede saber que es falsa, pero de cuya influencia no puede escapar. Por tanto, incluso si fuera verdad (que no lo es) que toda acción necesariamente tiene condiciones actitudinales entre sus causas precedentes, podrían, sin embargo, ser sólo fuerzas ajenas las que causaran que una persona lleve a cabo una acción.

La afirmación de que alguien realizó una acción implica que sus movimientos ocurrieron bajo su guía, pero no que la persona podría haberse abstenido de guiar sus movimientos como lo hizo. Hay ocasiones en las que actuamos contra nuestros deseos o independientemente de ellos. En otros casos, no estamos simplemente resignados al principio que guía nuestros movimientos, sino que,

más bien, lo hemos adoptado como propio. En dichos casos, por lo general tenemos una razón para adoptarlo. Quizá, como argumentarían algunos filósofos, el hecho de que tengamos una razón para actuar a veces puede hacer que los movimientos de nuestros cuerpos sean guiados por nosotros de un modo que refleje esa razón. Es indiscutible que las creencias y las actitudes de una persona a menudo tienen una relación importante con el modo como debe interpretarse y comprenderse lo que ella está haciendo; y es posible que a veces también aparezcan en las explicaciones causales de sus acciones. El hecho de que seamos racionales y conscientes afecta de manera sustancial el carácter de nuestra conducta y la manera en que nuestras acciones se integran en nuestras vidas.

V

No obstante, la importancia para *nuestras* acciones de los estados y los acontecimientos que dependen del ejercicio de nuestras capacidades superiores no debería llevarnos a exagerar la peculiaridad de lo que hacen los seres humanos. Estamos lejos de ser únicos tanto en el hecho de que nuestra conducta tenga un propósito como en su intencionalidad. Entre los filósofos, hay una tendencia a analizar la naturaleza de una acción como si la agencia presupusiera características que no pueden atribuirse en forma verosímil a miembros de otras especies que no sean la nuestra. Pero, de hecho, el contraste entre las acciones y los meros sucesos puede discernirse con facilidad en aspectos no vinculados con la vida de las personas. Hay numerosos agentes aparte de nosotros mismos que pueden ser activos o pasivos respecto de los movimientos de su cuerpo.

Consideremos la diferencia entre lo que sucede cuando una araña mueve sus patas al trasladarse por el suelo y lo que sucede cuando sus patas se mueven en patrones similares y con un efecto similar porque son manipuladas por un niño que ha logrado atarlas con hilos. En el primer caso, los movimientos no solamente tienen un propósito, como sin duda son los procesos digestivos de la

araña. También se le pueden atribuir a la araña que los crea. En el segundo caso, los mismos movimientos ocurren, pero no son creados por la araña, a quien meramente le suceden.

Este contraste entre dos tipos de acontecimientos en las vidas de las arañas, que pueden observarse en la historia de criaturas aun más ignorantes, es un paralelo del contraste, más conocido, entre el tipo de acontecimientos que ocurren cuando una persona levanta el brazo y el tipo de los que ocurren cuando su brazo se levanta sin que la persona lo mueva. Sin duda, se trata del mismo contraste en ambos casos. Las respectivas diferencias que los distinguen son similares, y tienen, por así decirlo, el mismo objetivo. Cada uno contrasta instancias en las que la conducta con propósito se le puede atribuir a una criatura como agente, e instancias en que ése no es el caso.

Este contraste genérico no puede explicarse en términos de ninguna de las facultades superiores distintivas que habitualmente entran en juego cuando una persona actúa. Evidentemente, las condiciones para atribuir la guía de los movimientos corporales a una criatura en su totalidad y no sólo a algún mecanismo local dentro de la criatura imperan fuera de la vida humana. De ahí que no puedan comprenderse en forma satisfactoria sobre la base de conceptos que no pueden aplicarse a las arañas ni a ninguna otra criatura de su especie. Esto no significa que deba ser ilegítimo para un análisis de la agencia humana invocar conceptos de alcance más limitado. Si bien las condiciones generales de una agencia no son claras, bien podría ser que el cumplimiento de estas condiciones por parte de seres humanos dependiera de la ocurrencia de acontecimientos o de estados que no ocurren en las historias de otras criaturas. Sin embargo, debemos cuidar que la manera en que interpretamos la agencia y definimos su naturaleza no esconda un sesgo limitado, que nos lleve a hacer caso omiso del grado en el cual el concepto de acción humana no es más que un caso especial de otro concepto, cuyo alcance es más amplio.

7
La importancia de lo que nos preocupa

I

Desde hace tiempo, los filósofos dedican una atención sistemática a dos amplios conjuntos principales de preguntas, que surgen a raíz de una preocupación con un aspecto omnipresente, imperioso y problemático de nuestras vidas. En el primer conjunto, que constituye el ámbito de la epistemología, las preguntas derivan de una u otra manera de nuestro interés en decidir *qué creer*. El tema general de las preguntas del segundo conjunto es *cómo comportarnos*, es decir, el contenido de la ética. También es posible delinear una tercera rama de investigación, que tiene que ver con una serie de cuestiones que pertenecen a otra área temática y son un objeto de interés fundamental de la existencia humana, es decir, *qué debe preocuparnos*.

Las diversas cuestiones conceptuales distintivas a las que conduce esta tercera preocupación no entran exactamente dentro del alcance de la investigación de la epistemología ni de la ética. Estas disciplinas no necesitan reflexionar sobre la naturaleza de la preocupación como tal, ni tampoco están obligadas a considerar lo que implica el hecho de que seamos criaturas a quienes las cosas nos importan. No trataré de proporcionar una descripción formal y exhaustiva de la rama de la investigación que se ocupa específicamente de esas cosas. En este ensayo, propongo sólo abordar, de un modo algo provisorio y fragmentario, algunos de sus conceptos y cuestiones centrales.

II

Naturalmente, existe una conexión muy estrecha entre lo que a una persona le preocupa y lo que –en general o en ciertas condiciones– piensa que es mejor hacer. Sin embargo, mientras que, por tanto, la tercera rama de la investigación se asemeja a la ética en su preocupación por los problemas de la evaluación y de la acción, difiere significativamente de la ética en sus conceptos generativos y en las preocupaciones que la motivan. La ética se concentra en el problema de ordenar nuestras relaciones con *otras personas*. Se ocupa especialmente del contraste entre lo *correcto* y lo *incorrecto*, y del fundamento y los límites de la *obligación moral*. Por otro lado, llegamos a la tercera rama de la investigación porque nos interesa decidir qué hacer con *nosotros mismos* y porque, en consecuencia, debemos entender qué es *importante* o, más bien, qué es *importante para nosotros*.

No se puede discutir que para la mayoría de nosotros los requerimientos de la ética no son las únicas cosas que nos preocupan. Hasta las personas a quienes les preocupa mucho la moralidad tienen en general otras preocupaciones mayores. Por ejemplo, tal vez les preocupen más sus propios proyectos personales, ciertos individuos y grupos y, quizá, diversos ideales a los que les otorgan una considerable autoridad en sus vidas, pero que no necesariamente deben ser de naturaleza ética. No hay nada distintivamente moral, por ejemplo, en ideales tales como ser tenazmente leal a una tradición familiar, perseguir desinteresadamente una verdad matemática o dedicarse a algún tipo de habilidad.

La función del juicio moral en el desarrollo y la búsqueda de preocupaciones como éstas suele ser bastante marginal, no sólo en potencia sino también en relevancia. De más está decir que hay muchas decisiones importantes con respecto a las cuales las consideraciones morales simplemente no son decisivas y que, como consecuencia, deben basarse, por lo menos en cierta medida, en consideraciones no morales. Sin embargo, incluso las decisiones que no son de esta clase muchas veces se toman, por supuesto, a la luz de valores o preferencias que no son morales. Más

aun, no es del todo evidente que tomarlas de esa manera sea siempre injustificado.

Alguien que se toma la moralidad en serio y que cree que una de sus alternativas es, en efecto, moralmente preferible a las demás, puede, no obstante, considerar que la importancia de este hecho no llega a ser categóricamente prioritaria. Supongamos, en primer lugar, que esa persona en realidad no sabe cuál de sus alternativas es la mejor desde el punto de vista moral. Podría ser sensato de su parte que rechazara de plano estudiar el asunto, con el argumento de que, dadas las circunstancias, hacerlo resultaría demasiado costoso. Es decir, podría juzgar razonablemente que sería más importante para él reservar para otros usos el tiempo y el esfuerzo que requerirían una exploración y una evaluación a conciencia de las características morales relevantes de su situación. La cuestión de si alguna vez un juicio de esta clase está completamente garantizado depende del hecho de si las consideraciones morales son necesariamente tanto más importantes que otras como para que no haya límites para los recursos que es razonable emplear a fin de asegurarse de que sean tenidas en cuenta.

O bien, en segundo lugar, supongamos que la persona ya sabe lo que está moralmente obligada a hacer. No obstante, puede elegir, de modo deliberado, violar esta obligación, no porque crea que ésta es anulada por otra más fuerte, sino porque existe un curso de acción alternativo que considera más importante para ella que cumplir con las exigencias de la rectitud moral. Me parece que tanto en este caso como en el primero podría justificarse la subordinación de unas consideraciones morales a las otras. Sea como fuere, en ambos casos está claro que la pregunta sobre qué es lo más importante puede distinguirse de la pregunta sobre lo que es moralmente correcto.

Puede haber personas para quienes las consideraciones éticas no sólo son indudablemente importantes, sino también exclusivas. Si así es, nada más tiene importancia alguna en sus vidas. Su único propósito, al que desean que contribuyan todas sus actividades, es hacer lo que consideran más deseable desde el punto de vista de la moralidad: maximizar el bienestar humano, quizá, o lograr que la sociedad sea más justa. Esta clase de sobreespecializa-

ción es difícil de sostener y es rara. No obstante, supongamos que alguien, en efecto, no acepte ninguna razón para actuar excepto que la acción en cuestión tenga más probabilidades que ninguna otra de conducir a la concreción de su ideal moral. Aun así, una cosa son los juicios morales de esta persona y otra, el hecho de que le preocupen tanto. En otras palabras, su creencia de que ciertas acciones están dictadas por consideraciones éticas difiere de su creencia de que ninguna otra consideración se compara en importancia con aquéllas.

III

Proporcionar análisis absolutamente inteligibles de los conceptos de preocupación y de importancia no es más fácil que definir las nociones –por ejemplo, la de creencia y la de obligación– básicas para las primeras dos ramas de la investigación. Sin duda, el concepto de importancia parece ser tan fundamental que quizá resulte absolutamente imposible hacer un análisis satisfactorio de él. Es razonable suponer que las cosas tienen importancia sólo en virtud de la influencia que tienen: si el hecho de que algo existiera o de que tuviera ciertas características no tuviera ninguna influencia sobre nada, entonces ni la existencia de esa cosa ni sus características tendrían importancia alguna. Sin embargo, en realidad todo tiene *alguna* influencia. ¿Cómo es posible, entonces, que algo en verdad carezca de importancia? Sólo puede ser porque la influencia de esa cosa no tenga importancia en sí misma. Así, en el análisis del concepto de importancia, evidentemente es fundamental incluir una condición a los efectos de establecer que nada es importante a menos que la influencia que tenga sea importante. No está claro si es posible elaborar una caracterización útil del concepto sin toparse con esta circularidad.

En cuanto a la noción de lo que le preocupa a la persona, en parte coincide con la noción de lo que le sirve de referencia a la persona para guiarse en lo que hace con su vida y en su conducta. Desde luego, no debe suponerse que toda vez que la vida de una

persona exhibe durante un período de tiempo una disposición más o menos estable de actitud o de conducta, ello refleje lo que le preocupa a la persona durante ese tiempo. Después de todo, los patrones de interés o de respuesta pueden ser sólo manifestaciones de hábitos o de regularidades involuntarias de otro tipo; y también es posible que aparezcan simplemente por casualidad. Por tanto, pueden ser discernibles incluso en las vidas de criaturas incapaces de preocuparse por algo.

El hecho de preocuparse, en la medida en que consista en guiarse por un curso distintivo o de una manera particular, presupone tanto acción como autoconciencia. Supone estar activo de algún modo, y con una actividad, en esencia, reflexiva. Esto no es exactamente porque el agente, al guiar su propio comportamiento, necesariamente se haga algo *a* sí mismo. Es más porque hace algo *con*sigo mismo con un fin determinado.

Puede decirse que una persona a la que algo le preocupa está investida en ello. Se *identifica* con aquello que le preocupa en el sentido de que se vuelve vulnerable a pérdidas y susceptible de beneficios según si lo que le preocupa disminuye o aumenta. De esta manera, se preocupa por lo que se relaciona con ello, presta particular atención a esas cosas y dirige su comportamiento en consecuencia. En la medida en que la vida de la persona esté *dedicada* por completo o en parte a algo, en lugar de ser sólo una secuencia de acontecimientos cuyos temas y estructuras no se esfuerza por moldear, está dedicada a esto.

Una persona podría dejar de preocuparse por algo porque sabe que no puede tenerlo. Sin embargo, podría seguir sintiéndose atraída por ello y deseándolo, y considerarlo deseable y valioso. Por tanto, el hecho de que algo nos preocupe no debe confundirse con que nos guste o que lo deseemos; tampoco es lo mismo que pensar que lo que nos preocupa tiene algún tipo de valor o que es deseable. Debe advertirse especialmente que estas actitudes y creencias difieren en forma significativa del hecho de preocuparse en sus características temporales. El punto de vista de una persona que se preocupa por algo es inherentemente prospectivo; es decir, la persona necesariamente considera que tiene un futuro. Por otro lado, es posible que una criatura tenga deseos

y creencias sin tomar en cuenta en absoluto el hecho de que puede seguir existiendo.

Los deseos y las creencias pueden aparecer en una vida que consiste simplemente en una sucesión de momentos separados, ninguno de los cuales el sujeto reconoce –ya sea cuando ocurre, ya en forma anticipada o en un recuerdo– como un elemento integrado con otros de su propia historia continua. Cuando este reconocimiento está complemente ausente, no *existe* un sujeto continuo. Se supone que las vidas de algunos animales son así. Sin embargo, los momentos en la vida de una persona que se preocupa por algo no están sólo vinculados inherentemente por relaciones formales de secuencialidad. La persona necesariamente los une y en la naturaleza del caso también interpreta que están unidos, de un modo más rico. Esto implica su propia preocupación continua por lo que hace consigo misma y con lo que sucede en su vida, y es implicado por ella.

Consideraciones similares indican que a una persona puede preocuparle algo sólo durante un período de tiempo más o menos extendido. Es posible desear algo o pensar que es valioso sólo por un momento. Los deseos y las creencias no tienen una persistencia inherente; no hay nada en la naturaleza del deseo o de la creencia que requiera que un deseo o una creencia deban perdurar. Sin embargo, la noción de guía y, de allí, la noción de preocupación, suponen cierta coherencia o firmeza del comportamiento; y esto presupone cierto grado de persistencia. Una persona a la que le preocupó algo sólo un momento no sería distinguible de alguien que fue inducido por un impulso. No estaría, en un sentido apropiado, guiándose o dirigiéndose a sí misma.

Dado que tomar una decisión requiere sólo un momento, el hecho de que una persona decida preocuparse por algo no puede ser equivalente a su preocupación por esa cosa. Tampoco es garantía de que ello efectivamente le preocupe. Al tomar esa decisión, la persona forma una intención con respecto a lo que debe preocuparle. Sin embargo, otra cosa muy diferente es si esa intención verdaderamente se satisface. Una decisión de preocuparse no implica más preocupación de lo que una decisión de dejar de fumar implica el dejar de hacerlo. En ninguno de los dos casos tomar la

decisión significa, siquiera, iniciar la situación que se ha decidido crear a menos que esa situación se produzca realmente.

No valdría la pena destacar esto, salvo porque en ocasiones se adjudica una importancia exagerada a las decisiones, a las elecciones y a otros *actos de la voluntad* similares. Si consideramos que la voluntad de una persona es aquella por la que dicha persona se mueve, entonces lo que le preocupa es más cercano al carácter de su voluntad que las decisiones que toma o las elecciones que hace. Las últimas pueden pertenecer a lo que la persona *tiene la intención* de que sea su voluntad, pero no necesariamente a lo que su voluntad realmente *es*.

A veces, se interpreta que el joven del famoso ejemplo de Sartre resolvió su dilema con respecto a si permanecer en su casa y cuidar a su madre, o abandonarla y sumarse a la lucha contra los enemigos de su país, mediante una elección radicalmente libre. Sin embargo, ¿cuán importante es el hecho de que el joven elija una de sus alternativas en lugar de la otra, incluso si comprendemos que su elección implica una decisión de su parte con respecto a qué clase de persona ser y no simplemente con respecto a qué hacer? Por cierto, no nos brinda ninguna razón en particular para pensar que en realidad él se convertirá en la clase de persona que decida ser, ni tampoco nos da derecho siquiera de suponer que en realidad llevará a cabo la alternativa que elija.

La cuestión no es que podría cambiar de parecer un momento después de tomar su decisión ni que podría olvidarla de inmediato. Es que podría ser *incapaz* de llevar a cabo su intención. Podría descubrir, a la hora de la verdad, que simplemente no puede realizar la acción por la que se ha decidido. Sin cambiar de parecer ni olvidar nada, podría descubrir o bien que se siente irresistiblemente inducido a seguir el otro curso de acción o bien que es obligado, de forma similar, por lo menos a abstenerse del curso de acción que eligió. O, en cambio, podría descubrir que en realidad es capaz de llevar a cabo las acciones que ha elegido realizar, pero sólo obligándose a hacerlo contra inclinaciones naturales poderosas y persistentes. Es decir, podría descubrir que no tiene y, en consecuencia, no desarrolla los sentimientos, actitudes e intereses constitutivos de la clase de persona que su decisión lo ha comprometido a ser.

La resolución del dilema del joven, entonces, no requiere sólo que decida qué hacer. Requiere que realmente se preocupe más por una de las alternativas a las que se enfrenta que por la otra; y requiere, además, que comprenda cuál de esas alternativas es la que en verdad le preocupa más. La dificultad en la que se encuentra se debe o bien a que no sabe cuál de las alternativas le preocupa más o bien a que ambas le preocupan por igual. Está claro que en ningún caso su dificultad será superada en forma confiable tomando una decisión.

El hecho de que alguien se preocupe por algo está constituido por un conjunto complejo de disposiciones y estados cognitivos, afectivos y volitivos. En ocasiones, puede ser posible que una persona, cuando hace una elección o toma una decisión, efectivamente logre preocuparse por algo o que algo le preocupe más que otra cosa. Sin embargo, eso depende de condiciones que no siempre prevalecen. Ciertamente, no puede suponerse que lo que a una persona le preocupa está en general bajo su control voluntario e inmediato.

IV

Por supuesto, existen amplias variaciones en el grado de fortaleza y de persistencia con que a las personas les preocupan las cosas. También es posible discriminar diferentes *formas* de preocupación, que no pueden reducirse de ningún modo obvio a diferencias de grado. Las más notables de ellas son, quizá, las numerosas variedades de amor. Otra distinción importante –que se relaciona, sin ser idéntica a ella, con la de si la preocupación puede ser iniciada por un acto voluntario– tiene que ver con la noción de si una persona puede evitar preocuparse como lo hace o no. Cuando una persona se preocupa por algo, puede depender completamente de ella que se preocupe por eso y que se preocupe tanto como lo hace. Sin embargo, en ciertos casos la persona es susceptible de una clase de necesidad conocida aunque algo oscura, en virtud de la cual su preocupación no está por completo bajo su control.

Hay ocasiones en que una persona se da cuenta de que lo que la preocupa le importa no sólo tanto sino de tal manera, que le es imposible abstenerse de un determinado curso de acción. Es de suponer, por ejemplo, que Luther hizo su famosa declaración, "Heme aquí; *no puedo hacer otra cosa*", en una ocasión semejante. Un encuentro de esta clase con una necesidad afecta típicamente a una persona no tanto empujándola hacia cierto curso de acción como haciendo que, de alguna manera, le resulte obvio que todas las alternativas evidentes para ese curso de acción son impensables. Estos encuentros difieren de las situaciones en las que una persona descubre que es incapaz de abstenerse de hacer algo, ya sea que lo quiera o no, porque se siente impulsada a actuar por un deseo o por alguna compulsión demasiado poderosos de vencer. También difieren de las situaciones en las que a la persona le queda claro que debe rechazar la posibilidad de abstenerse de una acción, porque tiene una muy buena razón para rechazarla; por ejemplo, porque abstenerse le resulta una opción demasiado poco atractiva o poco deseable.

Por otro lado, los encuentros con la necesidad de esta clase son, en ciertos aspectos, similares a situaciones como éstas. Se asemejan a los del último tipo –es decir, una persona no puede abstenerse de algo porque sus razones para no hacerlo son demasiado buenas– en el hecho de que la incapacidad de abstenerse no es una simple cuestión de capacidad deficiente por parte del agente. Se asemejan a los de la primera variedad –es decir, la persona es impulsada por una pasión irresistible o algo similar– en el hecho de que el agente siente que no tiene más opción que acceder a la fuerza por la que es obligado, aunque piense que sería mejor no hacerlo.

Por supuesto, está claro que la imposibilidad a la que Luther se refería no era una cuestión ni de lógica ni de necesidad causal. Después de todo, sabía muy bien que, en un sentido, era bastante capaz de hacer precisamente lo que decía que no podía hacer; es decir, tenía la capacidad de hacerlo. Lo que era incapaz de dominar no era el *poder* de abstenerse, sino la *voluntad*. Usaré el término *capacidad volitiva* para referirme a la obligación del tipo al que él declaró estar sometido. En la medida en que dicha obligación realmente imposibilite que una persona actúe de cualquier manera

salvo como actúa, lo imposibilita impidiéndole que *haga uso* de sus propias capacidades. Quizás haya un sentido en que Luther, aun si su declaración era verdadera, podría haber tenido la fortaleza suficiente para superar la fuerza que le impedía seguir cualquier curso de acción excepto el que siguió. Sin embargo, no pudo *lograr* superar esa fuerza.

Una persona que está sometida a la necesidad volitiva encuentra que *debe* actuar como lo hace. Por esta razón, puede parecer apropiado considerar que las situaciones que involucran la necesidad volitiva proporcionan ejemplos de pasividad. Sin embargo, la persona que se encuentra en una situación de esta clase no suele interpretar que el hecho de estar sometida a la necesidad volitiva implica que sea pasiva en absoluto. En general, las personas están muy lejos de considerar que la necesidad volitiva las vuelve observadoras indefensas de su propio comportamiento. De hecho, hasta pueden llegar a considerar que ella, en realidad, incrementa tanto su autonomía como su fuerza de voluntad.

Si una persona que es forzada por la necesidad volitiva es, por esa razón, incapaz de seguir un curso de acción determinado, ello no se debe a que simplemente sea demasiado débil para superar la obligación. Esa clase de explicación puede dar cuenta de la experiencia de un adicto, que se disocia de la adicción que lo fuerza, pero que fracasa en su intento de oponer sus propias energías al ímpetu de su hábito. No obstante, una persona que es forzada por la necesidad volitiva se encuentra en una situación muy diferente de aquélla. A diferencia del adicto, no accede a la fuerza que la obliga porque carezca de fuerza de voluntad suficiente para derrotarla. Accede a ella porque *no está dispuesta* a oponerse a ella y porque, además, su renuencia es, *en sí misma*, algo que no está dispuesta a alterar.

No sólo le preocupa seguir el curso de acción específico que es obligada a seguir. También le preocupa que le preocupe. Por tanto, trata de no sentirse críticamente afectada por cualquier cosa –en el mundo externo o dentro de sí– que podría distraerla o disuadirla tanto de seguir ese curso de acción como de preocuparse como lo hace por seguirlo. No logra superar la obligación a la que está sometida porque, en otras palabras, en realidad no quiere hacerlo. El conflicto del adicto contra su voluntad es que hay algo que real-

mente quiere hacer, pero que no puede hacer a causa de una fuerza que no es su propia voluntad y es superior a ella. En el caso de la persona obligada por la necesidad volitiva, también hay algo que no puede hacer, pero sólo porque en realidad no quiere hacerlo.

La razón por la que una persona no experimenta la fuerza de la necesidad volitiva como algo ajeno o externo a ella es, entonces, porque coincide con deseos –y, sin duda, en parte está constituida por ellos– que no son simplemente suyos, sino que se identifica con ellos en forma activa. Más aun, la necesidad, en cierta medida, se autoimpone. Es generada cuando alguien se exige evitar ser guiado en lo que hace por fuerzas que no son las que él más profundamente desea que lo guíen. A fin de evitar preocuparse por cualquier cosa tanto como se preocupa por ellas, se reprime o se disocia de toda motivación o deseo que considera incoherente con la estabilidad y la efectividad de su compromiso. De esta manera, la necesidad volitiva puede tener un efecto liberador: cuando alguien tiende a distraerse de la preocupación por lo que más le preocupa, la fuerza de la necesidad volitiva puede obligarlo a hacer lo que realmente quiere hacer.

Sin embargo, cualquiera sea la pertinencia y la validez de estas consideraciones, no explican cómo es posible que una persona sea obligada por una necesidad que ella misma se impone. Con seguridad, las personas muchas veces se obligan a actuar de cierto modo; por ejemplo, cuando se sienten fuertemente tentadas a actuar de otra manera. Pero el agotador empleo de la fuerza de voluntad en casos de ese tipo es complemente voluntario. El agente puede interrumpirlo cada vez que lo desee. Por otro lado, aunque la necesidad volitiva se autoimponga, debe de haber algún aspecto en el que se impone o mantiene en forma involuntaria.

La condición de que sea autoimpuesta ayuda a explicar el hecho de que sea liberadora en lugar de coercitiva; por ejemplo, el hecho de que respalde la autonomía de la persona en lugar de oponerse a su voluntad o ser independiente de ella. Sin embargo, no puede suceder que la persona que se exige evitar guiarse a sí misma en cierto modo cumpla con la autoimposición de este requisito simplemente llevando a cabo un acto voluntario. Debe ser una característica esencial de la necesidad volitiva el que se imponga a

una persona contra su voluntad. De lo contrario, sería imposible explicar el hecho de que la persona no pueda librarse de ella a voluntad, es decir, el hecho de que sea una clase genuina de necesidad.

Puede parecer difícil entender cómo es posible que la necesidad volitiva pueda, al mismo tiempo, ser autoimpuesta y ser impuesta contra la propia voluntad, o cómo es posible evitar la conclusión de que un agente que es obligado por la necesidad volitiva deba ser simultáneamente activo y pasivo con respecto a la misma fuerza. La resolución de estas dificultades radica en reconocer que: (a) el hecho de que una persona se preocupe por algo es un hecho vinculado con su voluntad, (b) la voluntad de una persona no necesita estar bajo su propio control voluntario, y (c) su voluntad puede ser igualmente suya aunque no se preocupe como lo hace por su propia voluntad.

Por tanto, la necesidad volitiva puede ser autoimpuesta en virtud de ser impuesta por la propia voluntad de la persona y, al mismo tiempo, impuesta contra su voluntad, en virtud del hecho de que su voluntad no es lo que es por su propio acto voluntario. De modo similar, en esos casos, la falta de voluntad no implica la pasividad. Una persona es activa cuando hace lo que hace por su propia voluntad, a pesar de que su voluntad no esté dentro del alcance de su control voluntario. Al parecer, entonces, a menos que una persona se preocupe por ciertas cosas más allá de que elija o no hacerlo, no será susceptible de la liberación que puede proporcionar la necesidad volitiva.

V

La sugerencia de que una persona puede ser liberada, en algún sentido, accediendo a un poder que no está sometido a su control voluntario inmediato se encuentra entre los temas más antiguos y persistentes de nuestra tradición moral y religiosa. Ciertamente debe reflejar alguna característica estructural muy fundamental de nuestras vidas. Sin embargo, esta característica ha sido poco explorada. Como consecuencia, somos incapaces de dar explicacio-

nes satisfactoriamente perspicuas y cabales de ciertos hechos que son centrales para nuestra cultura y nuestra visión de nosotros mismos: en particular, que las dos capacidades humanas que más valoramos son la de la racionalidad y la del amor, y que estas capacidades son valoradas no sólo por su utilidad para permitirnos adaptarnos a nuestros ámbitos naturales y sociales, sino también porque se supone que ponen a nuestra disposición experiencias o estados de plenitud y de libertad especialmente valiosos. La idea de que la racionalidad y el amor nos permiten alcanzar la libertad debería desconcertarnos más de lo que lo hace, dado que ambas cosas requieren que una persona *se someta* a algo que está más allá de su control voluntario y que puede ser indiferente a sus deseos.

Cuando accedemos a ser inducidos por la lógica o por el amor, en general no lo hacemos con el sentimiento de impotencia desanimada. Por el contrario, normalmente experimentamos, en ambos casos —ya sea que sigamos a la razón, ya a nuestros corazones–, un sentimiento de liberación y de mejoría. ¿Cómo se puede explicar esta experiencia? Al parecer, ella surge del hecho de que cuando una persona percibe algo como racional o amado, responde con un sentimiento que tiende a la *abnegación*. Su atención no se concentra simplemente en el objeto; de alguna manera, se fija en el objeto o es atrapada por él. El objeto la cautiva. La persona es guiada por las características del objeto más que por las propias. En general, se siente vencida, siente que la dirección de sus propios pensamientos y voliciones ha sido reemplazada. ¿Cómo debemos entender la paradoja de que una persona pueda sentir liberación y mejoría porque la han atrapado, la han hecho cautiva y la han vencido? ¿Por qué nos sentimos completamente realizados y consideramos que estamos en nuestra mejor forma cuando –mediante la razón o el amor– nos hemos perdido o hemos escapado de nosotros mismos?[1]

1 También somos susceptibles de ser vencidos por la belleza y la grandiosidad, y nos encontramos con experiencias similares, aunque tal vez no idénticas, cuando nos perdemos en la emoción de un momento o en el trabajo. Estas experiencias también tienden a ser liberadoras. Por otro lado, las experiencias de mucho temor o de dolor proporcionan sentimientos análogos a la abnegación de la razón y del amor en las que la pérdida del yo no se interpreta en general como algo satisfactorio o liberador. Parece improbable que esto sea sólo porque estas

La racionalidad y el amor implican abnegación por igual. Difieren en que la primera también es esencialmente *impersonal*. Lo esencial de esta diferencia entre la racionalidad y el amor no consiste en que lo que una persona ama dependa en gran medida de sus propias características particulares, mientras que esas características no desempeñan ninguna función a la hora de determinar lo que ella considera que la razón requiere o permite. Es evidente que los juicios que una persona hace con respecto a la racionalidad no dependen menos que cualquier otro acontecimiento de su vida de características contingentes de su naturaleza y de sus circunstancias. Lo que hace que estos juicios sean impersonales es que lo que afirman no se limita a la persona que los hace; más bien, queda implícito que cualquiera que no esté de acuerdo con esas afirmaciones debe de estar equivocado. En cambio, una declaración de amor es una cuestión personal, porque la persona que la hace no se compromete con ella a suponer que cualquiera que no ame lo mismo que ella está equivocado.

Ahora bien, los juicios morales también son impersonales, y en este aspecto su fuerza se diferencia fundamentalmente de la de la necesidad volitiva. Aun cuando la necesidad volitiva surja en conexión con acciones que son exigidas o prohibidas por el deber, no se deriva de las convicciones morales de la persona como tales, sino de la forma en que la persona se preocupa por ciertas cosas. Si una madre que está tentada de abandonar a su hijo descubre que simplemente no puede hacerlo, es probable que no sea porque conoce su deber (ni siquiera porque se preocupa por él). Es más probable que sea porque se preocupa por su hijo y por ella misma como madre que porque reconozca que abandonar al niño sería moralmente incorrecto. Por tanto, la coherencia no requiere que ella suponga que la acción que no puede realizar les resulte igualmente imposible a todas las madres cuyas circunstancias son parecidas a la suya.[2]

---

experiencias son menos placenteras que las del amor y la racionalidad, pero no queda claro qué es lo que permite explicar la diferencia.

2 Aunque la *coherencia* no le exige suponer esto, ella y otros podrían suponerlo de todas formas por otros motivos. Aunque no sea moralmente obligatorio que las madres se preocupen profundamente por sus hijos, una madre que no lo hace

De la misma manera, una persona que descubre que no puede poner en peligro un ideal al que se ha consagrado, a pesar de su ansiedad con respecto al costo de permanecer leal a él, probablemente no esté siendo inducida en primera instancia por consideraciones morales objetivas, aun cuando el ideal en cuestión sea de una variedad distintivamente moral. Supongamos que el ideal de alguien es ser meticulosamente honesto al llevar adelante sus negocios. Por supuesto, todos estamos moralmente obligados a ser honestos, pero de ello no se deduce que todos tengamos el deber de perseguir la honestidad como ideal de vida; es decir, asignarle al hecho de perseguirlo la atención y el interés preferentes que implica el compromiso hacia un ideal. El descubrimiento de una persona de que le resulta volitivamente imposible descuidar uno de sus ideales no debe equipararse, entonces, con su reconocimiento de un requisito ético.

En especial con respecto a aquellos que amamos y con respecto a nuestros ideales, es probable que estemos unidos por necesidades que tienen menos que ver con nuestra adhesión a los principios de la moralidad que con la integridad o la coherencia de un tipo más personal. Estas necesidades nos impiden traicionar las cosas que más nos preocupan y con las cuales, por tanto, nos sentimos más identificados. En un sentido que un análisis estrictamente ético no puede aclarar, lo que las necesidades impiden que violemos no son nuestros deberes ni nuestras obligaciones, sino a nosotros mismos.

VI

La voluntad de una persona se forma, fundamentalmente, cuando comienza a preocuparse por ciertas cosas y cuando lo hace más por algunas de ellas que por otras. Aunque tal vez estos procesos

---

podría ser blanco de críticas, no porque su actitud viole un deber, sino, por ejemplo, porque se la considere *antinatural* o *vergonzosa*, y se acuse a la mujer de carecer de importantes cualidades humanas.

no estén por completo bajo el control de su voluntad, es posible con frecuencia que ésta los afecte. Por ello, como también porque la gente, en general, está interesada en saber qué pensar de sí misma, una persona puede preocuparse por lo que la preocupa. Esto lleva a cuestiones acerca de la evaluación y la justificación.

El hecho de que lo que preocupa a una persona sea una cuestión personal no implica que *cualquier cosa* valga. Aun puede ser posible distinguir entre cosas que en cierta medida son dignas de preocupación y cosas que no lo son. Como consecuencia, puede resultar útil investigar qué hace que algo sea digno de preocupación, es decir, qué condiciones deben cumplirse si algo ha de ser apropiado o digno como ideal o como objeto de amor, y cómo debe decidir una persona, entre las diversas cosas dignas de preocupación, por cuál preocuparse. Aunque algunas personas pueden preocuparse, justificadamente, por cosas diferentes o preocuparse en forma diferente por las mismas cosas, esto no significa, por cierto, que su amor y sus ideales no sean susceptibles de una crítica importante de cualquier clase o que no puedan encontrarse principios analíticos generales que posibiliten la diferenciación.[3]

Muchas veces las personas no se preocupan por ciertas cosas que son muy importantes para ellas. Después de todo, es posible que no puedan reconocer que esas cosas tienen tanta importancia. Sin embargo, si hay algo por lo que una persona sí se preocupa, entonces eso significa que es importante para ella. Esto no se debe a que la preocupación implique algún tipo de juicio infalible con respecto a la importancia de su objeto. Más bien es porque la preo-

3 Una versión del escepticismo con respecto a estas cuestiones es la opinión de que en realidad no existe nada que sea digno de preocupación. Cualesquiera sean los méritos de este punto de vista, es importante no confundirlo con la afirmación más radical de que nada tiene importancia para nosotros ni suponer que implica esta última afirmación. Una persona que se preocupa por algo incurre, como consecuencia, en ciertos costos vinculados con el esfuerzo que significa investirse a sí misma y con la vulnerabilidad frente a la desilusión y a otras pérdidas que ello impone. En virtud de estos costos, es posible que algo sea importante para una persona sin que sea lo suficientemente importante como para que valga el tiempo que esa persona invierte en preocuparse. La opinión de que nada es digno de preocupación, por tanto, implica sólo que nada tiene *suficiente* importancia para que preocuparse por ello sea razonable.

cupación por algo *vuelve* importante esa cosa para la persona que se preocupa por ella.

Por supuesto, necesariamente una persona que se preocupa por algo no es fríamente indiferente a ello. En otras palabras, lo que le suceda a la cosa debe tener alguna influencia en la persona que se preocupa por ella, y la influencia en sí misma debe ser importante para esa persona. Naturalmente, esto no significa que se preocupa por ella sólo porque la afecta de modo importante. Por el contrario, puede ser que la persona sea susceptible de ser afectada por esa cosa o a causa de ella sólo en virtud del hecho de que se preocupa por ella.

Esto sugiere que lo que les preocupa a las personas es necesariamente importante para ellas. El hecho de que alguien se preocupe por una cosa o una persona, o el hecho de que no se preocupe por ellas, tiene una gran influencia sobre él. Significa que es susceptible, o no, de ser afectado por diversas circunstancias de modos que él considera importantes. Por tanto, la pregunta de por qué preocuparse (que, debe interpretarse, incluye la pregunta de si vale la pena preocuparse por algo) necesariamente debe ser importante para él.

De todo esto no se desprende que necesariamente valga la pena que una persona se preocupe por la pregunta. Ésta puede no ser lo suficientemente importante para él. Sin embargo, lo que sí se desprende es que si *hay algo que sea* digno de preocupación, entonces la pregunta por lo que hay que preocuparse debe ser digna de preocupación. Difícilmente suceda que haya algo tan importante para una persona que valga la pena que se preocupe por ello y que, a la vez, no valga la pena que se preocupe acerca de si le preocupa o no esa cosa.

En cualquier caso, existen dos formas distintas (aunque compatibles) en las que algo puede ser importante para una persona. En primer lugar, la importancia que tiene para esa persona puede deberse a consideraciones completamente independientes del hecho de si le preocupa o no la cosa en cuestión. En segundo lugar, la cosa puede volverse importante para la persona sólo porque se preocupa por ella. Como consecuencia, hay dos motivos distintos por los que una persona que piensa que vale la pena preocuparse por cierta cosa podría intentar justificar su punto de vista. Podría afir-

mar que la cosa es importante para ella en forma independiente y que, por esta razón, es digna de preocupación. O bien podría sostener, sin suponer que la cosa ya era importante para ella, que está justificada al preocuparse por ella porque hacerlo es, en sí mismo, algo que es importante para ella.

Es natural que las personas quieran que las cosas por las que se preocupan coincidan, hasta cierto punto, con las que son importantes para ellas, independientemente o previamente. Así, una persona muchas veces comienza a preocuparse por algo cuando reconoce la capacidad que ello tiene de afectarla de modo importante, deja de preocuparse por ello cuando descubre que no tiene esa capacidad y se critica a sí misma por preocuparse demasiado o demasiado poco por cosas cuya importancia para sí juzgó mal. Sin embargo, cuando la importancia de cierta cosa para una persona se debe al hecho de que se preocupa por ella, ese hecho simplemente no puede proporcionar una medida útil del grado en que su preocupación por la cosa está justificada.

En esos casos, la pregunta crítica no puede ser si el objeto es lo suficientemente importante para la persona para justificar que se preocupe por él. Debe ser, en cambio, si la persona tiene justificación para *volver* la cosa importante para ella preocupándose por esa cosa. Ahora bien, la única forma de justificar esto es en términos de la importancia de la actividad de preocuparse. Es evidente que los diversos modos de estar interesados o dedicados y de amar *son* importantes para nosotros independientemente de cualquier capacidad previa de afectarnos que tuviera aquello por lo que nos preocupamos. Esto no se debe en particular al hecho de que preocuparnos por algo nos hace susceptibles de ciertas gratificaciones y desilusiones adicionales. Es principalmente porque sirve para conectarnos activamente con nuestras vidas en formas que son creativas de nosotros mismos y que nos exponen a posibilidades distintivas para la necesidad y la libertad.

Sería un grave error creer que la importancia que un objeto tiene para alguien no es completamente genuina a menos que sea independiente de su preocupación por ese objeto. Consideremos el hecho de que muchas de las personas por las que más nos preocupamos no nos afectarían en forma importante si no nos preocu-

páramos por ellas. Esto no significa que no sean genuinamente importantes para nosotros. En ciertos casos, ciertamente, puede parecer que algo carece de verdadera importancia para una persona a pesar de que ésta se preocupa por ello. Sin embargo, si la importancia del objeto en esos casos no es completamente genuina, no es porque derive del hecho de que la persona se preocupa por el objeto.

Supongamos, por ejemplo, que aquello por lo que una persona se preocupa es evitar romper un espejo. Sin duda está cometiendo un error de algún tipo al preocuparse por esto. Pero su error no es que se preocupe por algo que en realidad no es importante para ella. Más bien, su error consiste en preocuparse por algo que no es digno de preocupación y, por tanto, conferirle una importancia genuina. La razón por la que eso no es digno de preocupación parece clara: no es importante para la persona hacer que evitar romper un espejo sea importante para ella. Sin embargo, necesitamos entender mejor por qué esto es así; es decir, qué condiciones deben cumplirse si para nosotros debe ser importante hacer que algo sea importante para nosotros que, de otra manera, no tendría esa importancia.

Aun cuando la justificación de una preocupación se apoye en la importancia de la preocupación misma y no derive de la importancia previa de su objeto, la elección del objeto no es irrelevante ni arbitraria. De acuerdo con una doctrina teológica, el amor divino es, en efecto, otorgado sin tener en cuenta el carácter o el valor previo de sus objetos. La naturaleza de Dios, desde esta perspectiva, es amar, y, por tanto, Él ama todo independientemente de cualquier consideración extrínseca a Sí mismo. Su amor es enteramente arbitrario e inmotivado, absolutamente soberano y de ninguna manera está condicionado por el mérito de sus objetos.[4] Quizá sólo un ser omnipotente –para quien nada es importante previamente– tiene la posibilidad de amar con total libertad y sin condiciones ni restricciones de ningún tipo. En cualquier caso, la capacidad de amor totalmente incondicional no es de ninguna manera un componente esencial de nuestra naturaleza finita.

4 Véase Anders Nygren, *Agape and eros*, Nueva York, 1969, pp. 75-81, 91-95.

¿Qué es lo que hace más apropiado para una persona, entonces, hacer que un objeto sea más importante para ella que otro? Al parecer, debe ser el hecho de que para ella es *posible* preocuparse por una cosa y no por otra, o preocuparse por ella de una forma que es más importante para ella que la forma en la que le es posible preocuparse por la otra. Como consecuencia, cuando una persona hace que algo sea importante para ella, la situación se asemeja a una instancia de *ágape* divino, por lo menos en un aspecto. La persona no se preocupa por el objeto porque el valor de dicho objeto la obligue a hacerlo, sino que el valor de la actividad de preocuparse la obliga a elegir un objeto por el que será capaz de preocuparse.

8
De qué somos moralmente responsables

Se podría haber esperado que naturalmente se interpretara que la libertad de la *voluntad* de una persona radica en la cuestión de si lo que desea depende de ella. De hecho, por lo general se entiende que está relacionada con la cuestión de si lo que *hace* depende de ella. En otras palabras, se considera que la voluntad de alguien es libre en un momento dado sólo si en ese momento depende de ella hacer una cosa u otra. Cuando este concepto de libre albedrío se suma a la suposición de que el libre albedrío es una condición necesaria de la responsabilidad moral, el resultado es el Principio de las Posibilidades Alternativas (PPA): una persona es moralmente responsable de lo que hizo sólo si podría haber hecho otra cosa.

Para aquellos que aceptan el PPA, resulta importante la pregunta de si las personas alguna vez *pueden* hacer alguna otra cosa de lo que en realidad hacen. Los incompatibilistas sostienen que si la tesis determinista es verdadera, esto no es posible. Por otra parte, los compatibilistas insisten en que, incluso en un mundo determinista, una persona puede tener alternativas genuinas en el sentido que requiere el PPA. Desde mi punto de vista, el PPA es falso.[1] El hecho de que una persona no tenga alternativas impide que sea moralmente responsable cuando esto por sí mismo da cuenta de su comportamiento. Sin embargo, la falta de alternativas no contradice la responsabilidad moral cuando alguien actúa como lo hace por razones propias y no simplemente porque no tiene otra alter-

1 Defiendo esta postura en "Posibilidades alternativas y responsabilidad moral", capítulo 1 de este volumen.

nativa. Por consiguiente, no tiene una especial importancia, en lo que concierne a las atribuciones de responsabilidad moral, si la tesis determinista es verdadera o falsa, o si es compatible o incompatible con el libre albedrío según lo interpreta el PPA.

I

El atractivo del PPA podría deberle algo a la suposición de que es un corolario de la tesis kantiana de que *debería* supone *puede*.[2] Pero, en realidad, la relación entre la doctrina de Kant y el PPA no es tan estrecha como parece. Respecto de cualquier acción, la doctrina de Kant se relaciona con la capacidad de un agente de llevar a cabo *esa* acción. El PPA, en cambio, se refiere a su capacidad de hacer *algo más*. Asimismo, el enfoque de Kant deja abierta la posibilidad de que una persona que sólo tiene un curso de acción disponible esté cumpliendo una obligación cuando lo sigue y sea moralmente digna de elogio por hacerlo. En cambio, el PPA implica que esa persona no puede tener ningún mérito moral por lo que hace. Esto deja en claro que renunciar al PPA no requiere negar que *debería* supone *puede*, y que el PPA no es una consecuencia de la perspectiva de Kant.

Imaginar contraejemplos para el PPA no presenta dificultades. Sólo es necesario concebir circunstancias que tornen inevitable que una persona lleve a cabo alguna acción sin que sean su causa. Por ejemplo, digamos que una persona decide consumir cierta droga y que lo hace, con el único fin de gozar de la euforia que espera que la droga le cause. Ahora bien, supongamos además que la acción de consumir la droga habría sido causada, de todas maneras, por fuerzas que, de hecho, estaban inactivas, pero que habrían entrado en juego si ella, por sí misma, no hubiese decidido actuar como lo hizo. Digamos que, sin saberlo, la persona es adicta a la droga y, por tanto, habría sido llevada de manera irresistible a

2 Véase R. Cummins, "Could have done otherwise", en *The Personalist*, LX, Nº 4, octubre de 1979, pp. 411-414.

consumirla si no lo hubiera hecho libremente. Su adicción latente garantiza que ella no podría haber evitado decidir consumir la droga ni consumirla, pero no interviene para provocar su decisión ni su acto. A medida que se desarrolla la secuencia real de acontecimientos, todo sucede como si la persona no fuera en absoluto adicta. La adicción es claramente irrelevante en este caso respecto de la cuestión de si la persona es moralmente responsable de consumir la droga.

El elemento distintivamente convincente en este tipo de contraejemplos del PPA es cierto tipo de sobredeterminación que incluye un orden secuencial infalible según el cual un factor causalmente suficiente funciona de manera exclusiva como respaldo de otro. El orden asegura que cierto efecto será provocado por uno u otro de los dos factores causales, pero no por los dos a la vez. Así, el factor de respaldo puede no contribuir en absoluto a provocar el efecto cuya ocurrencia garantiza.

II

Peter van Inwagen argumenta enérgicamente que incluso si los contraejemplos de este tipo –parafraseándolo, me referiré a ellos como *contraejemplos del tipo F*– requieren que se deje a un lado el PPA, la disputa entre el compatibilismo y el incompatibilismo conserva su importancia.[3] Él considera que la suposición de que el PPA es falso no implica, como he sostenido, la irrelevancia de la relación entre el determinismo y el libre albedrío respecto de las cuestiones que hacen a la responsabilidad moral. "Incluso si el PPA es falso", dice,

> es cierto, *no obstante*, que a menos que el libre albedrío y el determinismo sean compatibles, el determinismo y la responsabi-

3 Peter van Inwagen, "Ability and responsibility", en *Philosophical Review*, LXXXVII, Nº 2, abril de 1978, pp. 201-224. En lo que sigue, los números entre paréntesis se refieren a las páginas de su ensayo.

> lidad moral son incompatibles. En consecuencia, aun si los argumentos de Frankfurt son sólidos, no privan al debate entre compatibilistas e incompatibilistas de su lugar central en la antigua controversia acerca del determinismo y la responsabilidad moral (p. 223).

A fin de respaldar esta posición, Van Inwagen formula tres principios, que considera "muy similares al PPA", pero que cree que se diferencian de él por ser inmunes al tipo de objeciones que debilitan el PPA (p. 203). Arguye que demostrar esta inmunidad sirve para restablecer la importancia de las posibilidades alternativas y, por tanto, de la cuestión del compatibilismo, para la teoría de la responsabilidad moral.

Van Inwagen denomina su primer principio *Principio de la Acción Posible*:

PAP: Una persona es moralmente responsable de no realizar un acto dado sólo si hubiera podido realizarlo (p. 204).

Este principio se refiere a los "hechos no realizados (cosas que dejamos sin hacer)" (p. 203). El segundo y el tercer principio, que Van Inwagen denomina *Principios de la Prevención Posible*, tienen que ver con "las consecuencias de lo que hicimos (o dejamos sin hacer)" (p. 203):

PPP1: Una persona es moralmente responsable de cierto acontecimiento (particular) sólo si podía haberlo evitado (p. 206).
y
PPP2: Una persona es moralmente responsable de una situación sólo si (esa situación prevalece y) ella podría haber evitado que prevaleciera (p. 210).

El PPA sólo se refiere a la responsabilidad moral de una persona por *lo que hizo*. De esta manera, la suposición de que el PPA es falso deja abierta la posibilidad de que pueda haber cosas *diferentes* de los elementos de su propia conducta –a saber, acontecimientos, situaciones o actos no realizados que son consecuencia de lo que hizo

o dejó sin hacer– de los cuales una persona no puede ser moralmente responsable a menos que su voluntad sea o haya sido libre.

Van Inwagen brinda una versión del incompatibilismo para cada uno de sus tres principios. Según la primera de ellas, el determinismo implica que cualquiera que no haya realizado un acto dado no podría haberlo realizado. La segunda y la tercera se suman a la afirmación de que el determinismo implica que no hay acontecimientos o situaciones tales –y, por tanto, consecuencias de lo que alguien hizo– que cualquiera podría haber impedido que ocurrieran o prevalecieran. Ahora bien, Van Inwagen está convencido de que el PAP, el PPP1 y el PPP2 son verdaderos. Por consiguiente, su postura es que si las versiones del incompatibilismo que corresponden con esos principios también son verdaderas, entonces "el determinismo implica que nadie fue nunca o nunca pudo haber sido responsable de ningún acontecimiento, situación o acto no realizados" (p. 222).

Creo que, pese a que Van Inwagen lo niega, su primer principio es, en realidad, vulnerable a los contraejemplos del tipo F. En cambio, bien podría ser que la misma estrategia nó funcionara contra sus otros dos principios. Sin embargo, si ello es así, se debe sólo a que el PPP1 y el PPP2 son irrelevantes para la relación entre el libre albedrío y el determinismo. Su inmunidad a los contraejemplos del tipo F no brinda, por consiguiente, respaldo a las conclusiones que Van Inwagen propone extraer acerca de cómo la teoría de la responsabilidad moral se ve afectada por consideraciones relacionadas con el determinismo y el libre albedrío.

III

En su análisis del PAP, Van Inwagen no tiene en cuenta la sospecha bastante natural de que el principio simplemente equivale al PPA. En cambio, procede directamente a examinar un contraejemplo putativo. Debido a que considera que el ejemplo adapta la estrategia general que trae problemas al PPA al caso de los actos no realizados, supone que el ejemplo ofrece una prueba crítica del hecho de si esta estrategia es efectiva contra el PAP.

En el ejemplo, Van Inwagen es testigo de un crimen y piensa en llamar a la policía para denunciarlo. Debido a que no quiere verse involucrado, decide no llamar a la policía y no hace nada. No obstante, sin que él lo sepa, el sistema telefónico ha dejado de funcionar, y no se puede usar ninguno de los teléfonos a su alcance. Respecto de esta situación, plantea la siguiente pregunta: "¿Soy responsable de no haber llamado a la policía?". Su respuesta es inequívoca y enfática: "Por supuesto que no. No podría haberlo hecho". Dadas las circunstancias, dice: "Puedo ser responsable de no haber *tratado* de llamar a la policía (eso *podría* haberlo hecho), o de haberme abstenido de llamarla, o de [...] ser egoísta y cobarde. Pero simplemente no soy responsable de no haber llamado a la policía" (p. 205).

Entonces, según la opinión de Van Inwagen, el ejemplo deja al PAP totalmente ileso. Su conclusión es que el "estilo de argumentación de Frankfurt no puede emplearse para refutar el PAP" (p. 205).

Ahora bien, ser responsable de algo puede significar, en un sentido estricto de la noción, ser *completamente* responsable de ello, es decir, asignarle una condición *tanto* suficiente *como* necesaria. En este caso, una persona es completamente responsable de todos aquellos acontecimientos o situaciones, y sólo de ellos, que ocurren como resultado de lo que hace y que no ocurrirían si hiciera otra cosa. La persona del ejemplo de Van Inwagen (de ahora en más, *P*) no es, en este sentido, responsable de no haber llamado a la policía. Su comportamiento efectivo fue una condición suficiente para no hacerlo, pero no fue una condición necesaria: dado que el sistema telefónico no funcionaba, no habría llamado a la policía independientemente de lo que hiciera. Quizá, la razón por la que a Van Inwagen le parece tan obvio que la incapacidad de *P* de hacer lo que no hizo implica que *P* no es moralmente responsable de no haberlo hecho es que interpreta la responsabilidad moral en este sentido estricto de responsabilidad.

En mi opinión, la responsabilidad total no es una condición necesaria de la responsabilidad moral. Por tanto, creo que *P* puede ser moralmente responsable de no haber llamado a la policía aun cuando no hubiera podido evitar no haberlo hecho. Sin embargo, aquí no pretendo defender esta posición. En cambio, intentaré es-

tablecer otras dos cuestiones, que en mi opinión están más relacionadas con las conclusiones finales de Van Inwagen. La primera es que la pregunta acerca de si la responsabilidad moral presupone la responsabilidad total carece de interés *moral*. La segunda es que incluso si la responsabilidad moral presupone la responsabilidad total, el contraejemplo putativo que Van Inwagen presenta para el PAP no cumple su propósito.

IV

Supongamos que da la casualidad de que no sabemos si los teléfonos funcionaban cuando *P* tomó la decisión de no llamar a la policía y actuó en consecuencia. El hecho de que carezcamos de esta información no sería un obstáculo para que hiciéramos una evaluación moral competente de *P* por lo que hizo. Cuanto mucho, nos provocaría incertidumbre acerca de cómo *describir* el hecho de que *P* no haya actuado: es decir, cómo *identificar* acerca de *qué* lo estamos evaluando. Si los teléfonos funcionaban, sería más apropiado referirse al hecho de que no *llamó* a la policía; si no funcionaban, sería más apropiado referirse al hecho de que no *intentó llamar a la policía*. No obstante, la calidad del juicio moral y su grado –si *P* es digno de culpa o elogio, y en qué medida– serán exactamente los mismos en ambos casos.

Después de todo, cuál sea la inacción *de* la que se interprete que *P* es moralmente responsable depende por completo de las condiciones del sistema telefónico. De ninguna manera depende de alguno de sus actos, omisiones, estados psicológicos o propiedades, sean incorrectos o meritorios. La diferencia entre evaluar a *P* por no llamar y evaluarlo por no intentar llamar puede, por consiguiente, no tener importancia *moral*. Sólo es pertinente para decidir si es más adecuado expresar con unos términos o con otros lo que, en todo caso, continúa siendo el mismo juicio moral.

No se trata de que el comportamiento de una persona sea pertinente para los juicios morales que le conciernen sólo como prueba de su carácter o de su estado mental. Con seguridad, las intencio-

nes de *P* y sus otras características psicológicas continúan siendo idénticas, independientemente de si las circunstancias dictan que sea juzgado por no haber llamado a la policía o por no haber intentado llamarla. Pero la razón por la cual la evaluación moral de *P* será la misma en ambos casos no es que merece el elogio o la culpa moral exclusivamente por sus características psicológicas. Lo que una persona hace no es pertinente para las evaluaciones morales que recibe simplemente porque sea un indicador de su estado mental. La gente merece el elogio y la culpa *por* lo que hace y no sólo *sobre la base de* lo que hace.

Notemos que las intenciones de *P* y las demás cuestiones relacionadas no son las únicas cosas que permanecen iguales si los teléfonos funcionan o no. También está claro que, cualquiera sea el estado del sistema telefónico, *P* hace o no hace los mismos movimientos corporales. Ahora bien, *de lo que P es moralmente responsable* es de hacer estos movimientos. Es moralmente responsable de hacerlos, por supuesto, sólo en ciertas condiciones; sólo, por ejemplo, cuando los hace con ciertas intenciones o expectativas. Sin embargo, si se satisfacen esas condiciones, sólo es moralmente responsable de realizar los movimientos en sí mismos.

Hay diversas maneras en que se pueden identificar o describir los movimientos de una persona. El hecho de si es más apropiado definir lo que hace *P* como llamar a la policía o sólo como tratar de llamarla dependerá mucho de las consecuencias que tengan sus movimientos. Y las consecuencias de sus movimientos dependerán, a su vez, de si los teléfonos funcionan. No obstante, la calidad y el grado de la responsabilidad moral de *P* por lo que hace continúan siendo los mismos en ambos casos precisamente porque se juzga a *P* sólo por realizar sus movimientos. No hace falta decir que sus movimientos no son afectados por las consecuencias que provocan, sin que importe con cuánta determinación esas consecuencias pueden afectar los términos mediante los cuales resulta apropiado describir los movimientos.[4]

4 Evoco aquí la conocida interpretación de Donald Davidson, elaborada con persuasiva lucidez en su ensayo titulado "Agency", en Brinkley, Bronaugh y Marras (comps.), *Agent, action and reason,* Nueva York, Oxford, 1971, pp. 3-25, según la cual "nunca hacemos más que mover nuestro cuerpo; el resto depende de la

V

Debido a que *P* no habría llamado a la policía independientemente de lo que hiciera, no es *completamente* responsable de no haberla llamado. Esto le ofrece a Van Inwagen una razón para afirmar que *P* no es *moralmente* responsable por no haberlo hecho. Pero incluso si suponemos que la responsabilidad moral exige responsabilidad en el sentido estricto, el juicio de que el PAP es inmune a los contraejemplos tampoco tendría justificación. Ello es porque la razón por la que el contraejemplo de Van Inwagen no es efectivo reside en las características específicas del hecho de que *P* no haya llamado a la policía y no en las características de cualquier falta de acción. Por consiguiente, el contraejemplo no ofrece una prueba decisiva para el PAP.

¿Por qué *P* no es completamente responsable de no haber llamado a la policía? Porque no tiene el poder de hacer que el teléfono de la policía suene: independientemente de los movimientos corporales que haga, sus movimientos no tendrán las consecuencias que deben ocurrir para que sea correcto afirmar que *P* llamó a la policía. Sin embargo, también hay casos en los que la acción no se realiza que, a diferencia del de *P*, no dependen en absoluto de las consecuencias de lo que una persona haga. Son acciones no realizadas, de las cuales los movimientos mismos de una persona, considerados por completo aparte de sus consecuencias, son una condición tanto suficiente como necesaria. Por ejemplo, supongamos que *Q* está conduciendo su automóvil y no puede mantener la vista al frente porque prefiere mirar el interesante panorama que tiene a su izquierda; y supongamos, también, que si el panorama no lo hubiese distraído, alguna otra cosa habría causado que estuviera mirando hacia su izquierda en ese momento. En estas circunstancias, *Q* no puede mantener la vista hacia el frente. ¿Es moralmente responsable de no hacerlo? ¡Por supuesto que lo es! El

naturaleza" (p. 23). Si adaptamos y parafraseamos su versión (véase p. 21) al caso de *P*, podría decirse que después de que *P* movió sus manos de la manera en que deben moverse para hacer un llamado telefónico, hizo su trabajo; sólo faltaba que la compañía telefónica hiciera su parte.

hecho de que no pueda evitar no hacerlo no tiene relación con su responsabilidad moral por no haberlo hecho, puesto que no lo conduce a no hacerlo.

Notemos que *Q* es *completamente* responsable de no mantener la vista hacia adelante. No hacerlo depende, exclusivamente, de los movimientos que realiza una persona; está *constituido* por lo que la persona misma hace, y lo que la persona hace es, por consiguiente, una condición tanto suficiente como necesaria para ello. Entonces, no puede decirse que el hecho de que *Q* no mantuviera la vista hacia adelante habría sucedido independientemente de lo que hiciera, es decir, independientemente de los movimientos corporales que hizo. De no haber movido sus ojos hacia la izquierda, no habría dejado de mantener la vista hacia adelante. Así, la razón para negar que *Q* es moralmente responsable de no haberlo hecho no es la misma que para negar que *P* es moralmente responsable de no haber llamado a la policía. Incluso si se admite que la responsabilidad moral presupone una responsabilidad total, es posible encontrar contraejemplos efectivos contra el PAP.

Evaluar el PPA y los tres principios que presenta Van Inwagen es cuestión de decidir si una persona puede ser moralmente responsable de llevar a cabo una acción o de no hacerlo, o de las consecuencias de lo que hizo, a pesar de que la acción, el no llevarla a cabo o sus consecuencias no podrían haberse evitado. Ahora bien, hay dos maneras en que la acción de una persona, su falta de acción o una consecuencia de lo que ha hecho pueden ser inevitables. Pueden ser inevitables en virtud de ciertos movimientos que hace la persona y que no puede evitar hacer, o pueden ser inevitables debido a acontecimientos o situaciones que con seguridad van a ocurrir o prevalecer, independientemente de lo que haga la persona. A falta de una mejor terminología, me referiré al primer tipo de inevitabilidad como *personal,* y al segundo como *impersonal.*

Aparentemente, Van Inwagen supone que una persona no puede ser completamente responsable de no hacer algo que es incapaz de evitar. Esta suposición sería correcta si todas las faltas de acción inevitables fueran como la falta de acción de *P,* es decir, si su inevitabilidad fuera siempre *impersonal.* Sin embargo, en realidad el carácter inevitable de algunas faltas de acción es *personal.* En cuanto

a estos casos, no es verdad que ocurrirán independientemente de lo que haga la persona en cuestión. Son inevitables precisamente porque la persona, como Q, no puede evitar hacer los movimientos corporales que las constituyen.

VI

Creo que no hay una diferencia inherente entre realizar una acción y no hacerlo en virtud de la cual el PAP podría ser verdad aun cuando el PPA fuera falso. El PAP tampoco es inmune al tipo de contraejemplos ante los cuales sucumbe el PPA. Por otra parte, hay una variante del PAP que *sí* goza de inmunidad ante los contraejemplos del estilo F. Esta versión restringida del PAP, que denominaré *PAP'*, se refiere exclusivamente a las faltas de acción cuya inevitabilidad es impersonal. Se refiere a la responsabilidad moral de una persona por no provocar algún acontecimiento o situación cuando el hecho de que no lo haga es independiente de lo que la persona misma hace; es decir, de los movimientos que realiza. Así el PAP' se asemeja mucho al PPP_1 y al PPP_2, dado que ellos también se refieren a la responsabilidad moral de una persona por acontecimientos o situaciones que pueden ocurrir o prevalecer independientemente de los movimientos que la persona realice.[5]

Ahora bien, es verdad, por definición, que una persona no puede ser *completamente* responsable de algo que sucede o acontece inde-

5 Como ya he señalado, Van Inwagen dice que el PPP_1 y el PPP_2 tienen que ver con "las consecuencias de lo que hicimos (o dejamos de hacer)" (p. 203). Esto no está explicitado en su formulación de los principios. De hecho, son vulnerables, en una lectura natural, a cualquier contraejemplo efectivo contra el PPA. Ello se debe a que, si suponemos que hacer algo implica que ocurra un acontecimiento y prevalezca una situación, cualquier cosa que muestre que una persona puede ser moralmente responsable de lo que hizo, incluso si no podría haber hecho otra cosa, también muestra que puede ser moralmente responsable de ese acontecimiento o situación que ella no podría haber evitado. Sin embargo, en vista de la afirmación de Van Inwagen de que el PPP_1 y el PPP_2 se refieren a las *consecuencias* de lo que hace la gente, interpretaré que *no* se refieren a los movimientos corporales que hace la gente o a lo que esos movimientos necesariamente implican.

pendientemente de sus propios movimientos corporales. El hecho de si puede ser *moralmente* responsable de esas cosas depende de la relación entre la responsabilidad moral y la responsabilidad total. Supongamos que la primera presupone la última. En ese caso, el PAP', el PPP1 y el PPP2 son inmunes a los contraejemplos del estilo F. No obstante, esto no supone, como evidentemente cree Van Inwagen, que hay cosas de las cuales una persona puede ser moralmente responsable sólo si su voluntad es libre. A diferencia del PPA y el PAP, los tres principios en cuestión no tienen nada que ver con el libre albedrío.

El hecho de que haya acontecimientos o situaciones que una persona no puede causar claramente no significa, en sí mismo, que la persona carezca de libre albedrío. Aceptando que la libertad de la voluntad de una persona estriba esencialmente en la cuestión de si lo que hace depende de ella, se trata más de si depende de ella qué movimientos corporales hace que de las consecuencias que pueden provocar sus movimientos. Imaginemos que el mal funcionamiento del equipo que imposibilita que *P* llame a la policía, a pesar de su libertad para mover su cuerpo como lo desee, se debe a la negligencia de la compañía telefónica, e imaginemos que debido a esta negligencia una gran cantidad de personas está impedida de hacer diversas cosas. Es posible que estas personas se sientan, con razón, bastante enojadas. Pero estarían llevando su enojo demasiado lejos y le estarían atribuyendo a la compañía telefónica un papel demasiado significativo en sus vidas si se quejaran de que la negligencia de la compañía les cercenó la libertad de su voluntad.

De la misma manera que el PAP', el PPP1 y el PPP2 no tienen nada que ver con la relación entre la responsabilidad moral y el libre albedrío, tampoco tienen nada que ver con la relación entre el libre albedrío y el determinismo. Supongamos que no está determinado causalmente si los teléfonos funcionan. Entonces, tampoco está determinado si *P* puede llamar a la policía. Sin embargo, esto no implica nada en relación con la libertad de la voluntad de *P*.

Desde luego, tiene una relación significativa con la medida de su *poder*; es decir, con la efectividad de lo que hace. Sin embargo, no afecta de ninguna manera su libertad de moverse como quiere, ni es pertinente para la pregunta de si sus movimientos mismos

carecen de determinación. En otras palabras, es totalmente irrelevante para los tipos de intereses y angustias por los cuales las personas han sido empujadas a resistir la doctrina de que la vida humana está total e inevitablemente sujeta a la determinación causal.

9
Necesidad y deseo

I

El lenguaje de la *necesidad* se usa en forma extendida en la representación de nuestra vida personal y social. Su papel en el discurso político y moral es especialmente conspicuo y poderoso. La gente se suele atribuir necesidades y se las atribuye a los demás a fin de respaldar exigencias, establecer derechos o ejercer influencia en la jerarquía de prioridades; y, con frecuencia, tendemos a responder a dichas atribuciones con un respeto y un interés bastante especiales. En particular, la afirmación de que algo es *necesario* tiende a crear la impresión de una cualidad completamente diferente y a causar un impacto moral sustancialmente mayor que la afirmación de que algo es *deseado*. Las aseveraciones basadas en lo que una persona necesita suelen tener un patetismo característico. Es probable que susciten un sentimiento más apremiante de obligación y que sean tratadas con mayor urgencia que aseveraciones basadas, simplemente, en lo que alguien quiere.

Sin embargo, hay que tener cuidado de no exagerar la superioridad inherente de las afirmaciones basadas en necesidades sobre las afirmaciones basadas en deseos. Ciertamente no ocurre que la fuerza moral de las necesidades sea incondicionalmente mayor que la de los deseos en el sentido de que a toda necesidad, sin excepción, deba asignársele una prioridad incondicional sobre cualquier deseo. Existen muchas ocasiones en las que es perfectamente sensato que una persona sacrifique algo que necesita, aun algo que necesita mucho, para obtener algo que desea, pero

que no necesita en absoluto. Por ejemplo, sería muy sensato que una persona gravemente enferma usara sus limitados recursos económicos para salir en el crucero de placer que siempre deseó en vez de usarlos en la cirugía que necesita para prolongar su vida. Las decisiones de disfrutar más de la vida a costa de no cuidar de nosotros mismos como podríamos hacerlo –de mejorar la calidad de vida a costa de su cantidad– no son ni infrecuentes ni siempre injustificadas.

Quizás esto sea insuficiente para mostrar que una aseveración basada en el deseo puede competir con éxito, por motivos *morales*, con una aseveración respaldada por la necesidad. Sin embargo, de hecho las necesidades pueden no ser más apremiantes que los deseos aun respecto de consideraciones estrictamente morales. Consideremos a una persona que desea completar un crucigrama y que es incapaz de hacerlo sin buscar las palabras en el diccionario. Necesita un diccionario, pero la importancia moral de su necesidad es completamente insignificante. No sería difícil encontrar numerosos deseos con por lo menos la misma importancia moral.

Sin embargo, ahora parece que si una necesidad puede ser completamente intrascendente, las atribuciones de la necesidad en realidad no tienen, después de todo, ningún peso moral inherente. Este resultado parece ser decididamente confirmado, además, por consideraciones teóricas elementales. Nada es necesario excepto para un fin para el cual es indispensable. La importancia moral de satisfacer o no una necesidad, por tanto, debe derivar por completo de la importancia del fin que da origen a aquélla. Cualquiera sea la importancia de lograr ese fin, será igualmente importante satisfacer la necesidad. Si la importancia moral de la necesidad de un diccionario es insignificante, se debe a que el objetivo del que deriva la necesidad no tiene consecuencias morales. Así, parece que la satisfacción de necesidades no tiene derecho a ninguna prioridad moral sistemática sobre la gratificación de los deseos. El solo hecho de que algo sea necesario, considerado en forma aislada del valor de aquello para lo que se lo necesita, no tiene ningún peso independiente para justificarlo.

Sin embargo, debemos tener cuidado tanto para evitar exigir demasiado poco para las necesidades como para evitar exigir dema-

siado para ellas. Incluso fuera de otras consideraciones, la idea de que el hecho de que una persona necesite algo no tiene ninguna importancia moral especial es difícil de conciliar con la manifiesta potencia retórica de ciertos usos vagamente manipuladores a los que el lenguaje de la necesidad es sometido con frecuencia. Éstos suelen desdibujar la distinción entre necesitar algo y quererlo, con la obvia intención de atraer para algún deseo el mismo grado de consideración moral que tiende a ser adjudicado a las necesidades en particular.[1]

Las maniobras de esta clase no tendrían sentido a menos que las personas estuvieran ampliamente dispuestas a aceptar la proposición de que una necesidad de algo anticipa un deseo *de esa cosa.* Esta proposición, que llamaré *el Principio de Precedencia,* les atribuye a las necesidades sólo una mínima superioridad moral sobre los deseos. Sólo sostiene que cuando existe una competencia entre un deseo y una necesidad por la misma cosa, la necesidad comienza con cierta ventaja moral. Es decir, cuando A necesita algo que B quiere, pero que no necesita, entonces satisfacer la necesidad

1 Jean-Paul Sartre y Fidel Castro colaboran en la siguiente conversación para producir un egregio ejemplo:
"—La necesidad del hombre es su derecho fundamental sobre todos los demás –dijo Castro.
—¿Y si a usted le piden la Luna? –preguntó Sartre.
—... Sería porque alguien la necesita –fue la respuesta de Castro" (citado por George Lichtheim en *The concept of ideology and other essays,* Nueva York, 1967, p. 282).
Ahora bien, del hecho de que alguien pida algo a lo sumo se entiende, por supuesto, que lo quiere. Esta clase de confusión entre lo que se quiere y lo que se necesita es bastante común entre los marxistas. De esta manera, aunque parece obvio que algunas mercancías pueden satisfacer sólo deseos, el propio Marx define una mercancía como "una cosa que merced a sus propiedades satisface necesidades humanas del tipo que fueran" (*El capital,* Moscú, 1961 [citamos la traducción al español de México, Siglo XXI, 1987 (16ª ed.), t. I, vol. I, p. 43]). Daré por sentado que querer algo no implica necesitarlo y viceversa: una persona puede desear someterse a una cirugía sin necesitar una operación en realidad, o puede necesitar una cirugía sin querer someterse a ella. Esto no implica que los conceptos de necesidad y de deseo sean lógicamente independientes. Serían lógicamente independientes si y sólo si algo pudiera ser deseado sin que nada fuera necesario y algo pudiera ser necesario sin que nada fuera deseado. Lo que muestra el ejemplo es sólo que alguien puede tener una necesidad sin tener un deseo *de lo que necesita,* y que puede tener un deseo sin tener necesidad *de lo que quiere.*

de A es, *prima facie*, moralmente preferible a satisfacer el deseo de B. Si las necesidades no gozan por lo menos de *esta* precedencia sobre los deseos, entonces ciertamente debe de ser un error atribuirles *cualquier* importancia moral. En cualquier caso, el Principio parece ser sumamente razonable. Si no intervienen otros factores, parece claro que es preferible asignar un recurso a alguien que lo necesita más que asignárselo a alguien que lo quiere, pero que no lo necesita en absoluto.

Sin embargo, hay excepciones al Principio de Precedencia. Supongamos que alguien emprende cierto proyecto sólo por un capricho irreflexivo. El hecho de que, de ahí en más, necesite lo que sea indispensable para llevar a cabo el proyecto no tiene más fuerza justificativa de la que tendría un deseo casual o impulsivo por la misma cosa. La demanda de una persona que necesita un diccionario sólo para gratificar su capricho de terminar un crucigrama no tiene más peso que la demanda de alguien que no tiene ninguna necesidad específica de un diccionario, pero cuyo deseo es poseer uno sin ninguna razón en particular. El hecho de darle precedencia a la necesidad en este caso asignaría arbitrariamente una mayor importancia moral a un capricho que al otro.

La importancia moral de una necesidad no es, entonces, necesariamente mayor que la del deseo correspondiente. Por tanto, está claro que no podemos aceptar la doctrina de que es moralmente preferible asignar recursos a quienes los necesitan más que a quienes sólo los desean. Debemos distinguir entre los tipos de necesidades que merecen precedencia sobre los deseos que les corresponden, y los tipos de necesidades que no la merecen.

II

En el centro del concepto de necesidad está la noción de que hay cosas de las que uno no puede prescindir. Cuando algo es necesario, por tanto, siempre debe ser posible especificar *para qué* se lo necesita o explicar *qué* es lo que no se puede hacer sin él. Si una persona necesita una cirugía para sobrevivir, lo que no puede ha-

cer sin la cirugía es seguir viviendo. Todas las necesidades, en este aspecto, son condicionales: nada es necesario excepto en virtud de que es una condición indispensable para el logro de un fin determinado.[2]

En muchos casos, una persona necesita algo porque activamente desea un fin determinado para cuyo logro esa cosa es indispensable.[3] Así, la persona de mi ejemplo necesita un diccionario porque quiere terminar un crucigrama. De hecho, necesita el diccionario *sólo* porque quiere terminar el crucigrama; no lo necesitaría de no ser por ese deseo. Pero, por supuesto, una persona puede necesitar ciertas cosas por más de una razón, o en más de una forma. Cuando una persona necesita algo porque hay otra cosa que *quiere*, en ese caso la necesidad depende de la *voluntad* de la persona. Me referiré a las necesidades de esa clase como *necesidades volitivas*.

Tener una necesidad volitiva no es necesariamente un asunto voluntario, ya que la voluntad de una persona no está invariablemente bajo su control voluntario. Es decir, el hecho de si tiene el deseo del que depende su necesidad volitiva puede no depender de ella. Muchos de los deseos de una persona, en efecto, son voluntarios, puesto que derivan simplemente de sus propias decisiones. Es

2 Al parecer, está implícito en el concepto de necesidad que aquello para lo que algo es necesario debe ser distinto de eso mismo. Por ello, resulta algo disonante sugerir que la vida y la felicidad se encuentran entre las cosas que las personas necesitan. Puede haber circunstancias en las que el que una persona siga viviendo o sea feliz, en realidad, sirva a un propósito especial; y en esos casos, puede ser apropiado decir que la persona necesita vivir o ser feliz. Sin embargo, no suponemos que el valor de la vida o de la felicidad deriva, en general, del valor de otra cosa.

3 Joseph Raz me ha señalado que una persona puede querer algo y, sin embargo, no necesitar ciertas cosas que son indispensables para su logro, porque está claro que sería incapaz de lograrlo aunque las tuviera. Si ella reconoce que no puede satisfacer ninguna serie de condiciones suficientes para el logro de lo que quiere, entonces no necesita satisfacer las condiciones necesarias. Se aplican consideraciones similares si, por alguna razón que no sea la imposibilidad de lograr aquello que quiere –por ejemplo, una baja prioridad–, la persona no espera ni tiene la intención siquiera de tratar de satisfacer su deseo. Cuando me refiero a lo que una persona "desea activamente", mi intención es excluir los deseos que no espera ni tiene intención de tratar de satisfacer. En adelante, supondré, sin explicitarlos como tales, que los deseos de los que se dice que dependen las necesidades son activos en este sentido.

común que alguien adquiera el deseo de ver una película determinada, por ejemplo, al elegir la película que va a ver. Los deseos de esta clase no se suscitan dentro de nosotros; se forman o se construyen mediante actos de voluntad que realizamos nosotros mismos, muchas veces muy alejados de cualquier estado emocional o afectivo. Sin embargo, también hay ocasiones en las que lo que una persona quiere no depende de ella, sino que más bien es una cuestión de sentimientos o inclinaciones que surgen y persisten independientemente de cualquier elección que haga.

Ahora bien, supongamos que, con respecto a cierto deseo, el tenerlo o no tenerlo depende de la persona. Entonces, también depende de ella el tener o no una necesidad volitiva de aquello que es indispensable para la satisfacción de ese deseo. En cambio, si no tiene control sobre lo que quiere, tampoco tiene control sobre la cuestión de si tiene o no necesidades volitivas de esas cosas sin las cuales el deseo en cuestión no puede satisfacerse. Me referiré a las necesidades volitivas que dependen de deseos voluntarios como *libres*, y a aquellas que dependen de deseos involuntarios como *forzadas*.

Las necesidades volitivas libres no son, como tales, moralmente interesantes en el sentido especificado por el Principio de Precedencia. En otras palabras, no merecen tener prioridad sobre los deseos que les corresponden. A partir del hecho de que una persona necesita M porque es indispensable para E, que es lo que él quiere, no podemos llegar a la conclusión de que la consideración a la que tiene derecho su necesidad de M es mayor que la consideración que merecería por un simple deseo de M. No hay razón para pensar que su demanda de M recibe más apoyo de su deseo de E que lo que recibiría la demanda de M de otra persona del deseo de M de esa persona. Después de todo, ¿por qué el segundo deseo debería ser portador de una demanda menor que el primero? El hecho de que una persona desee M mientras que otra persona tiene una necesidad volitiva libre de M deja abierta la pregunta sobre cuál de las demandas de M es mejor.

Si las necesidades volitivas libres, como tales, carecen de importancia moral, no es porque los deseos de los que derivan sean uniformemente insignificantes. El hecho de que un deseo sea volun-

tario no implica nada en absoluto con respecto a su importancia. Una persona puede decidir por su propio libre albedrío no sólo que quiere terminar un crucigrama, sino también asuntos mucho más trascendentes: que quiere ser músico, que quiere renunciar a sus obligaciones y dedicarse a perseguir sus despiadados intereses materiales, que quiere morir, y así sucesivamente. Los deseos de los que dependen las necesidades volitivas libres de una persona pueden cambiar su vida en forma considerable.

Si no intervienen otros factores, la conveniencia de satisfacer una necesidad volitiva libre depende por completo de cuán deseable sea satisfacer el deseo voluntario pertinente. Hasta el punto en que sea deseable satisfacer el deseo de alguien de un fin determinado, será deseable, en la misma medida, satisfacer las necesidades generadas por ese deseo. Así, la conveniencia del fin de una persona puede justificar su demanda de lo que necesita a fin de lograrlo. Sin embargo, en la medida en que su deseo de ese fin sea voluntario, la conveniencia de satisfacerlo no puede dotar a su demanda de la cualidad moral *distintiva* que es específica de las demandas justificadas por la necesidad.

Ello se debe a que las necesidades volitivas libres contienen demasiado poca necesidad. Existen aquí dos consideraciones relacionadas, que iluminan la precedencia moral sobre los deseos, de la que gozan las necesidades de cierto tipo. En primer lugar, dado que el deseo del que deriva una necesidad volitiva libre puede tener cualquier objeto, quizás no sea ni importante ni necesario satisfacer el deseo; por tanto, no puede suponerse que las necesidades de esta clase *necesitan satisfacerse*. En segundo lugar, a partir del hecho de que el deseo que genera una necesidad volitiva libre es voluntario, se desprende que la persona que tiene dicha necesidad no *necesita necesitar* lo que necesita. Por otro lado, a fin de ser interesante desde el punto de vista moral, una necesidad debe ser radicalmente distinta de un deseo. Debe ser lo que llamaré *categórica*; es decir, estar caracterizada por las dos necesidades que hemos considerado: (1) la necesidad debe ser una que la persona no sólo quiere satisfacer sino que necesita satisfacer, y (2) lo que la persona necesita debe ser algo que no puede evitar necesitar. Analizaré estas dos condiciones más adelante.

III

La razón por la que las necesidades volitivas libres no necesitan, como tales, ser satisfechas es que los deseos de los que dependen pueden ser de cosas que no son necesarias. En ese caso, la persona *quiere* que su necesidad sea satisfecha para poder gozar de lo que desea, pero no *necesita* que sea satisfecha más de lo que necesita la cosa deseada en sí. Supongamos que resulta que no puede satisfacer su necesidad volitiva libre y que, por tanto, no puede tener lo que quiere. Entonces, la persona bien puede sentirse decepcionada y a la vez tener motivos para estar enojada. Sin embargo, dado que lo que quiere no es algo que necesita, *no se habrá causado ningún daño*. La persona se habrá privado de un beneficio de mayor o menor valor, pero no habrá salido perjudicada.

La vinculación con el daño diferencia de otras las necesidades que satisfacen el Principio de Precedencia y que, por tanto, son moralmente interesantes. La necesidad de una persona tiene interés moral sólo si es una consecuencia de su incapacidad de satisfacer la necesidad o si continúa sufriendo algún daño. Por supuesto, esta condición puede cumplirse aunque la persona no desee el objeto que necesita. En la medida en que el vínculo con el daño no dependa de un deseo, la necesidad es no volitiva. Las necesidades volitivas libres no tienen ningún interés moral inherente porque el solo hecho de que una persona tenga cierto deseo indica, a lo sumo, que espera que lo que desea signifique algún beneficio para ella. No implica que sufrirá algún daño si no lo obtiene.

No está claro cómo distinguir en forma sistemática las circunstancias en virtud de las cuales una persona es dañada de aquéllas en virtud de las cuales simplemente deja de obtener un beneficio; tampoco es evidente cómo definir esas condiciones especiales en las que alguien que deja de obtener un beneficio de ese modo también sufre un daño. Una forma de tratar este último problema sería sostener que esa incapacidad de obtener un beneficio es equivalente a sufrir un daño sólo en caso de que el beneficio sea algo que la persona en cuestión necesita. Esto es razonable, pero por razones obvias no resulta muy útil en este contexto. En vez de tratar

de formular una explicación más satisfactoria del tema, me limitaré a tres observaciones elementales pertinentes a la relación entre los beneficios y los daños.

En primer lugar, cuando se sufre un daño se pasa a estar en una situación peor que la anterior, mientras que cuando no se obtiene un beneficio, lo que ocurre es más bien que no se logra una situación mejor que la anterior. En segundo lugar, a veces no hay forma de impedir que una situación empeore excepto mejorándola. En situaciones como ésta, el hecho de no obtener el beneficio pertinente equivale a ser perjudicado. En tercer lugar, la vida de una persona cuya situación es mala empeora cada vez más a menos que su situación mejore, simplemente porque más de algo malo es peor que menos de algo malo. Alguien puede resultar perjudicado, por tanto, aun cuando en cierto sentido su situación no se deteriore. Esto hace posible aprobar el juicio de sentido común de que un enfermo crónico tiene una necesidad moralmente relevante de cualquier tratamiento que sea indispensable para aliviar su enfermedad, puesto que esto implica que, aunque el estado de la salud de la persona siguiera siendo básicamente el mismo, no sólo no obtendría un beneficio valioso si no recibiera el tratamiento, sino que, en realidad, resultaría perjudicada.

Estas observaciones sugieren por qué satisfacer las necesidades merece recibir prioridad respecto de la satisfacción de los deseos. La razón de ello es que mejorar la situación es, desde un punto de vista moral, menos importante (a igual medida) que evitar que empeore. En general, cuando se confía algo al cuidado de una persona, esperamos que ésta haga un esfuerzo razonable para protegerlo del daño o el perjuicio; pero normalmente no suponemos que tenga una obligación comparable de mejorar su situación. Más generalmente, con respecto a esa parte del mundo que depende del cuidado de alguien –es decir, de la que es responsable–, su obligación de evitar que empeore es más imperiosa que su obligación (si es que existe) de mejorarla. Por ello, el hecho de asignar recursos a fin de satisfacer necesidades tiene precedencia sobre asignarlos para satisfacer meros deseos. Lo primero apunta a evitar el daño, mientras que lo segundo apunta sólo a proporcionar beneficios innecesarios.

Las necesidades moralmente interesantes de una persona deben satisfacerse, entonces, porque de lo contrario sobrevendrá un daño. Pero, además, el vínculo con el daño debe ser de una naturaleza tal que el hecho de que el daño ocurra o no esté fuera del control voluntario de la persona. Éste es el segundo aspecto en que las necesidades volitivas libres contienen demasiado poca necesidad. No sólo derivan de deseos, lo cual significa que puede no haber daño alguno aunque no se satisfagan, sino que, además, los deseos de los que derivan son voluntarios, lo cual significa que la persona no necesita tener las necesidades.

Supongamos que en virtud de su propia decisión con respecto a lo que quiere, una persona tiene el deseo del que deriva cierta necesidad. Esto no la pone en poder de la necesidad. El poder que la tiene prisionera es el suyo, del que puede liberarse como quiera. No es sorprendente que las necesidades de esta clase no provoquen ninguna preocupación moral en particular. Aun cuando la persona, en efecto, sufra algún daño si no obtiene el objeto que necesita, ésta es una consecuencia que ella se impone a sí misma y a la que sigue expuesta sólo mientras esté dispuesta a ello. Necesita el objeto, ya que es indispensable para alcanzar un fin que desea. Sin embargo, su necesidad de él es su propia creación. El carácter indispensable del objeto para lograr el fin concierne a la persona sólo mientras que ella así lo quiera. No la afecta a menos que, por su propia elección, adopte el deseo pertinente.

IV

Ni los deseos ni las necesidades volitivas libres están vinculados en forma ineludible con el daño. Por ello, no son moralmente distinguibles los unos de las otras y ambos difieren de las necesidades categóricas desde el punto de vista moral. En efecto, las necesidades volitivas libres no sólo no merecen precedencia sobre los deseos correspondientes, sino que tampoco existe una razón para atribuirles, como tales, ningún interés moral en absoluto. Es decir, ni siquiera podemos suponer que satisfacer necesidades

de esta índole es inherentemente deseable o preferible a no satisfacerlas.

El hecho de satisfacer necesidades volitivas libres sería inherentemente deseable sólo si fuera inherentemente deseable satisfacer deseos. Sólo en ese caso se podría suponer que es conveniente satisfacer cualquier necesidad volitiva libre dada. Ahora bien, algunos filósofos sostienen que es necesariamente deseable que un deseo sea satisfecho. Así, William James escribe: "Tomemos cualquier demanda, por ligera que sea, que una criatura, por débil que sea, pueda hacer. ¿Acaso no debería, sólo por su bien, ser satisfecha? [...] Todo deseo es imperativo en la medida de su magnitud; se *hace* válido por el hecho de existir".[4] Por supuesto, James reconocería que la conveniencia de satisfacer un deseo puede ser anulada por otras consideraciones. Sin embargo, en su opinión, el hecho de que una persona quiera algo es siempre una razón en sí misma para que sea preferible que lo tenga.

En cambio, en mi opinión, el solo hecho de que una persona quiera algo no proporciona ningún respaldo para afirmar que el hecho de que lo tenga es preferible al hecho de que no lo tenga. No es mi intención negar que es mejor que algunos de los deseos de una persona sean satisfechos antes que no satisfacer ninguno. Tal vez, si no intervienen otros factores, necesariamente es mejor que algunos de los deseos de la vida de una persona sean satisfechos en vez de que ninguno lo sea. Sin embargo, de esto no se desprende que con respecto a cada uno de los deseos de una persona, es mejor que tenga lo que quiere en lugar de que no lo tenga. Lo que se desprende es sólo que el hecho de que una persona tenga algunas de las cosas que quiere es preferible a que no tenga ninguna.

Hasta donde yo sé, el único argumento disponible para defender la postura a la que adhiere James es, más o menos, el siguiente: un deseo insatisfecho inevitablemente supone frustración, lo cual es desagradable. Por esta razón, siempre existe, por lo menos, la misma consideración en favor de satisfacer un deseo dado que la que existe en favor de minimizar la sensación desagradable. Ahora

4 William James, "The moral philosopher and the moral life", en *Essays in pragmatism by William James*, ed. de Alburey Castell, Nueva York, 1948, p. 73.

bien, en efecto, hay una suposición en favor de minimizar la sensación desagradable. Por tanto, siempre hay un argumento *prima facie* para satisfacer un deseo determinado más que para no satisfacerlo. Un deseo es "imperativo en la medida de su magnitud", como lo expresa James, porque la sensación desagradable, que es consecuencia de la frustración, será más o menos severa y, por tanto, más o menos indeseable, según cuán fuerte sea el deseo frustrado.

Sin embargo, lo máximo que puede inferirse válidamente de las premisas de este argumento es que existe un argumento *prima facie* contra la conveniencia de cualquier situación en la que alguien tiene un deseo insatisfecho. En otras palabras, la única suposición justificada es, simplemente, que los deseos satisfechos son preferibles a los frustrados.[5] Esto difiere sustancialmente de una suposición en favor de la satisfacción del deseo, porque un deseo satisfecho no es la única alternativa posible a uno frustrado. Después de todo, una persona también evita la frustración cuando –al ser persuadida o por alguna otra razón– se da por vencida o pierde su deseo sin satisfacerlo. Algunos de los métodos que pueden resultar efectivos para eliminar los deseos de una persona sin satisfacerlos son, ciertamente, muy objetables. No obstante, lo mismo puede decirse de algunos de los métodos mediante los cuales se pueden satisfacer los deseos.

La tesis de James reduce la distancia conceptual entre la necesidad y el deseo al vincular el deseo con el daño y, así, dar a entender que el hecho de querer algo significa necesitarlo. Si fuera inevitable que un deseo no satisfecho se frustrara, una persona no podría evitar el disgusto a menos que consiguiera lo que quería. Ahora bien, es razonable suponer que experimentar un disgusto equivale a sufrir un daño y que todo el mundo quiere evitarlo, de modo que todo el mundo necesita, tanto volitiva como no volitivamente, lo que sea indispensable para evitar el disgusto. Precisamente porque un objeto de deseo puede, en realidad, *no* ser indispensable para que alguien alcance esta meta, querer algo no implica necesitarlo. Dado que un deseo puede abandonarse o perderse, una persona puede ser capaz de evitar la frustración sin obtener

5 Véase Gary Watson, "Free agency", en *Journal of Philosophy*, Nº 72, 1975, pp. 210-211.

lo que quiere. Así, la satisfacción de un deseo no es necesariamente necesaria para evitar el daño.

V

Sin embargo, con respecto a algunas de las cosas que una persona quiere, puede no ser posible que logre dejar de quererlas o sea conducida a hacerlo. Esto no se debe a que los deseos en cuestión sean especialmente intensos o difíciles de controlar. Incluso los deseos moderados y fáciles de manejar pueden ser persistentes e imposibles de erradicar. Llamo *necesidades volitivas forzadas* a las necesidades generadas por deseos de esta clase, que deben ser satisfechos o frustrados. Está claro que suponen una necesidad mayor que las necesidades volitivas libres. Una persona cuya necesidad volitiva forzada no es satisfecha, sin importar lo que elija o haga en forma voluntaria, inevitablemente sufrirá algún daño; a saber, frustración. Esto basta para convertir dichas necesidades en categóricas y para justificar su gratificación por encima de la gratificación de los deseos que les corresponden.

Todas las necesidades volitivas forzadas satisfacen el Principio de Precedencia. Sin embargo, algunas de ellas sólo parecen dignas de un interés bastante restringido o equívoco. Lo que las distingue no es que los daños con los que están vinculadas sean relativamente poco importantes, puesto que los daños, en realidad, pueden ser muy graves. Más bien es que las necesidades parecen ser en cierto modo gratuitas o, incluso, perversas. Por ejemplo, supongamos que un hombre está obsesionado por la idea fija de que su vida no tendrá ningún valor a menos que él posea un determinado automóvil deportivo, y supongamos que la frustración de su deseo del automóvil sería tan profunda que, en efecto, le arruinaría la vida. El hombre no puede evitar querer el automóvil y lo quiere de manera tan intensa que experimentará un sufrimiento permanente y agobiante a menos que lo obtenga. Como aquí existe un vínculo con el daño sustancial, que no está bajo el control voluntario del hombre, su necesidad del automóvil deportivo es tanto categórica

como grave. ¿Cuál es la razón, entonces, para nuestra inquietud respecto de ella? ¿Por qué nos sentimos inclinados a no ser tan entusiastas para reconocer que la demanda que hace es realmente legítima?

Es probable que nuestra reacción frente a la necesidad que el hombre tiene del automóvil sea el resultado de una serie de consideraciones. Quiero centrar la atención en particular sobre una que no tiene nada que ver con ningún juicio con respecto a la mezquindad de su ambición o la superficialidad de su carácter. Sin duda, nuestro respeto por el hombre se ve perjudicado en forma significativa por nuestro sentimiento de que el objeto de su deseo no es digno de la enorme importancia que tiene para él. Sin embargo, nuestra respuesta a su necesidad también se ve afectada por una característica que esa necesidad comparte con otras cuyos objetos son mucho más dignos de deseo y preocupación que los automóviles deportivos: es decir, la necesidad del hombre tiene menos que ver con las características específicas de su objeto que con la naturaleza de su deseo de ese objeto.

El hecho de que el hombre sufra si no consigue el automóvil no se deberá a su velocidad o a su belleza, ni siquiera a su valor social. Es de suponer que quiere el automóvil en virtud de estas características, pero ellas no explican el hecho de que lo necesite. Hasta podríamos sugerir que lo que realmente necesita no es el automóvil como tal, sino la gratificación de su deseo. Su necesidad está ineludiblemente vinculada con el daño sólo en virtud de su deseo y no en virtud de las consecuencias que le traería prescindir del automóvil. Si no quisiera el automóvil deportivo como lo quiere, no lo necesitaría de un modo moralmente significativo. En otras palabras, no tiene ninguna necesidad no volitiva a la que corresponda su deseo del automóvil.

La cuestión puede aclararse si distinguimos las necesidades de esta clase de las necesidades que resultan de la adicción. Estas últimas comúnmente están asociadas con necesidades volitivas forzadas, pero ellas mismas no son volitivas en esencia. El adicto a la heroína en general tiene un deseo involuntario de la heroína; sin embargo, es más probable que este deseo surja a causa de su necesidad de la droga y no que la necesidad derive del deseo. En cual-

quier caso, ser adicto a algo no supone ser incapaz de evitar quererlo. El sufrimiento característico al que están sometidos los adictos a la heroína no es el dolor del deseo frustrado. Es un estado más específico, causado por la falta de heroína. Ocurre independientemente de lo que el adicto –que quizá no sabe a qué es adicto, o siquiera que haya algo a lo que es adicto– quiere o no quiere.

Existen dos tipos de situaciones en las que participan las necesidades volitivas forzadas. En las situaciones de un tipo, una persona tiene una necesidad no volitiva y también una necesidad volitiva forzada de cierto objeto; y, por tanto, necesitaría el objeto aunque no lo deseara. En las situaciones del otro tipo, la necesidad de la persona es exclusivamente volitiva; es decir, la persona necesita un objeto determinado sólo porque lo desea.

Dado que el adicto tiene una necesidad no volitiva de la heroína, su deseo involuntario de la droga sirve para un propósito útil. Lo induce a obtener algo que necesita y que no puede evitar necesitar independientemente de que lo desee. En cambio, el deseo (por ejemplo, el de un automóvil deportivo) del que depende la necesidad volitiva forzada de una persona no sirve para ese propósito cuando la persona no tiene una necesidad no volitiva que corresponda al deseo. En ese caso, no hay ninguna necesidad ni riesgo de daño fuera del deseo. El deseo no responde a una necesidad ni la refleja; la crea. Ahora bien, esta creación de un riesgo de dañar de ninguna manera mejora ni el valor inherente del objeto deseado ni su disponibilidad. Así, somete a una persona a cargas y riesgos adicionales sin brindarle ningún beneficio compensatorio. En este aspecto, las necesidades del tipo en cuestión son gratuitas o perversas.

VI

La gama y la gravedad de las necesidades de una persona dependen de lo que quiere, de cómo lo quiere y de esos aspectos no volitivos de su situación que determinan lo que la dañará y lo que la protegerá del daño. Esto significa que las necesidades pueden ser generadas, alteradas o eliminadas por cambios en el entorno y por el

curso natural de la vida humana. Además, las necesidades de los tres tipos que he considerado pueden verse afectadas por la acción humana deliberada o involuntaria.

Muchos críticos sociales sostienen que una de las formas en las que las sociedades explotadoras dañan a sus miembros es haciendo que sufran diversas necesidades que los críticos caracterizan como *falsas* o *inauténticas* o a las que se refieren de alguna otra manera que sugiere inconveniencia o defecto. Quizá podríamos preguntar si es deseable tener necesidades categóricas. La pregunta se refiere al hecho de si estaríamos mejor si no fuéramos vulnerables al daño o si, de alguna manera, es algo bueno que no seamos, en este aspecto, omnipotentes. En cualquier caso, quienes condenan la creación de necesidades falsas o inauténticas no tienen la intención de argumentar contra el aumento en la carga de necesidad que lleva la gente. Su queja es contra los aumentos de una clase más específica. Lo que consideran objetable en la creación de una necesidad falsa no es que se haya creado otra necesidad, sino que la necesidad que ha sido creada es falsa.

Sugiero que un criterio que captura por lo menos un elemento importante de lo que es objetable en ciertas necesidades –necesidades que es razonable considerar *falsas*– puede basarse en la diferencia entre las necesidades volitivas forzadas que coinciden con necesidades no volitivas y las que no lo hacen. Según este criterio, la necesidad de una persona de cierto objeto es *verdadera* o *auténtica* sólo si la persona necesita el objeto más allá del hecho de que lo quiera o no. En cambio, una necesidad es *falsa* o *inauténtica* si la persona necesita el objeto sólo porque lo desea. En otras palabras, las necesidades volitivas son verdaderas o auténticas sólo en la medida en que reflejan necesidades que son no volitivas.

Esta explicación atraviesa la distinción entre las necesidades que son naturales y las necesidades que son impuestas socialmente. Lo que hace que una necesidad sea falsa no es que tenga causas de cierto tipo. Las necesidades pueden ser auténticas o verdaderas aun cuando no sólo sean artificiales en el sentido de ser producidas por artimañas humanas, sino también cuando las artimañas sean maliciosas o injustas. La falsedad de una necesidad no se debe al hecho de haberse originado en la maquinación o en la negligencia de

los reaccionarios o de los malos, sino de que es gratuita o perversa de un modo que ya se ha indicado. Las necesidades falsas son aquellas en las que no hay más necesidad que la creada por el deseo. Su defecto es análogo al de las verdades de Protágoras, las cuales –de acuerdo con la representación de su doctrina que consta en el *Teeteto*– son creadas por completo a partir de creencias. Así como no es correcto interpretar una creencia como la medida de la verdad, tampoco es correcto considerar que el deseo puede ser la medida de la necesidad.

Existe una diferencia entre nuestra respuesta a las necesidades que surgen exclusivamente de la volición forzada y nuestra respuesta a las necesidades que no son en absoluto volitivas. Esta diferencia subsiste aun cuando, como en el caso de la adicción autoinducida, la necesidad no volitiva de alguien es el resultado de su propio comportamiento voluntario. Las necesidades que la naturaleza impone a una persona (aun cuando esto sea producto de su conducta) nos inclinan a sentir un interés con mayor compasión y empatía que las que derivan inmediatamente de la propia voluntad de la persona (incluso cuando ella no tiene control sobre lo que quiere). Nuestro sentimiento de que nos incumbe asistir a una persona necesitada tiende a atenuarse de alguna manera cuando la necesidad deriva esencialmente del deseo de esa persona.

Esto puede deberse a que la transformación del deseo en necesidad nos parece una analogía de *mala fe*, de modo que sospechamos que la persona en cuestión es incapaz de controlar su deseo sólo porque en realidad no quiere hacerlo. En ese caso, no consideramos que la necesidad sea completamente forzada y, por tanto, no la interpretamos como genuinamente categórica. Es posible que también haya otra razón. Al tratar de evitar el daño al que la expone una necesidad volitiva forzada, una persona está luchando no tanto con la naturaleza como consigo misma. Tal vez esto atenúe nuestro sentimiento de camaradería hacia ella. Si ella estuviera luchando contra la naturaleza –que es nuestro enemigo común–, nuestro instinto de aliarnos con ella sería más apremiante.

10
Sobre el concepto de *bullshit**

Una de las características más destacadas de nuestra cultura es la presencia de gran cantidad de lo que llamamos *bullshit*. Todo el mundo lo sabe. Cada uno de nosotros contribuye a ello. Sin embargo, tendemos a dar por sentada la situación. La mayoría de las personas confían bastante en su capacidad de reconocer lo que es *bullshit* y evitar ser engañadas. Por consiguiente, el fenómeno no ha suscitado gran interés reflexivo ni ha dado origen a un estudio continuado.

Como consecuencia, no nos queda claro qué es el *bullshit*, por qué hay tanto ni qué función cumple. Y no hemos elaborado a conciencia una apreciación de lo que significa para nosotros. En otras palabras, no tenemos ninguna teoría. Propongo que intentemos llegar a una conceptualización teórica del *bullshit* ofreciendo, principalmente, un análisis filosófico tentativo y exploratorio. No tendré en cuenta los usos erróneos ni retóricos del *bullshit*. Mi objetivo es simplemente brindar una explicación aproximada de qué es lo que llamamos *bullshit* y en qué se diferencia de lo que no lo es, o (en otras palabras) expresar, a grandes rasgos, la estructura de su concepto.

Cualquier sugerencia acerca de qué condiciones son, desde un punto de vista lógico, necesarias y suficientes para la constitución

* La palabra *bullshit* se mantendrá en el idioma original, ya que no es posible encontrar un equivalente exacto en español. Si bien *bullshit* está compuesta por las palabras *bull* –que significa "toro"– y *shit* –"mierda"–, se podría traducir, en forma aproximada, como "tonterías", "mentiras", "sandeces", "joder", etc., según el contexto. [N. de T.]

del *bullshit* debe ser algo arbitraria. Por un lado, la expresión *bullshit* se suele utilizar con bastante libertad, simplemente como un término genérico referido al abuso, sin un significado literal muy específico. Por otro, el fenómeno en sí mismo es tan amplio y amorfo que cualquier análisis claro y preciso de su concepto resultaría inflexible. Sin embargo, debería ser posible decir algo útil al respecto, aunque difícilmente sea decisivo. Después de todo, aun las preguntas más básicas y preliminares acerca del *bullshit* permanecen no sólo sin ser respondidas, sino incluso sin ser formuladas.

Por lo que sé, se ha trabajado muy poco en este tema. No he emprendido una revisión de la bibliografía, en parte porque no sé cómo abordarla. Por cierto, un lugar bastante obvio donde buscar es el *Oxford English Dictionary*. El *OED* presenta una entrada para el término *bullshit* en los volúmenes complementarios, y también contiene entradas para diversos usos pertinentes de la palabra *bull* y algunos otros términos afines. Consideraré algunas de estas entradas a su debido tiempo. No he consultado diccionarios de otras lenguas que no sean el inglés, porque no conozco los equivalentes de *bullshit* o de *bull* en ningún otro idioma.

Otra fuente que vale la pena es el ensayo que da el título a la obra *The prevalence of humbug*, de Max Black.[1] No estoy muy seguro de cuánto se acercan los significados de la palabra *humbug** al de la palabra *bullshit*. Por supuesto, ambas palabras no se pueden intercambiar con libertad una por la otra; está claro que se las emplea de manera diferente. Sin embargo, en general, parece que la diferencia tiene más que ver con consideraciones de cortesía y algunos otros parámetros retóricos que con los modos de significación estrictamente literales que más me interesan. Es más cortés, y también menos intenso, decir *Humbug!* que *Bullshit!* Para este análisis, supondré que no existe ninguna otra diferencia importante entre ambas.

Black sugiere varios sinónimos de *humbug*, entre ellos: *balderdash* ("tonterías"), *claptrap* ("paparruchadas"), *hokum* ("patrañas"), *drivel* ("estupideces"), *buncombe* ("charlatanería"), *imposture* ("impostura"), *quackery* ("fraude"). Esta lista de curiosos equiva-

1 M. Black, *The prevalence of humbug*, Ithaca, Cornell University Press, 1985.

* Una traducción aproximada sería *farsa*. [N. de T.]

lentes no resulta muy útil. Sin embargo, Black también enfrenta el problema de establecer la naturaleza del *humbug* en forma más directa, ya que ofrece la siguiente definición formal:

> *HUMBUG*: tergiversación engañosa, que no llega a ser una mentira, en especial mediante una palabra o un hecho pretencioso, de los pensamientos, los sentimientos o las actitudes de alguien.[2]

Probablemente pueda ofrecerse una formulación muy similar que enuncie las características esenciales del *bullshit*. Antes de desarrollar mi propia consideración de esas características, comentaré los diversos elementos de la definición de Black.

Tergiversación engañosa: la frase puede sonar pleonástica. Sin duda, lo que Black tiene en mente es que el *humbug* necesariamente pretende engañar, que su tergiversación no es sólo involuntaria. En otras palabras, es una tergiversación *deliberada*. Desde el punto de vista de la necesidad conceptual, si la intención de engañar es una característica invariable del *humbug*, la cualidad de *humbug* depende, al menos en parte, del estado de ánimo del perpetrador. Como consecuencia, no coincide con ninguna propiedad –ya sea inherente o relacional– que pertenezca únicamente al enunciado mediante el cual se perpetra el *humbug*. En este aspecto, la cualidad de *humbug* se parece a la de la mentira, que no coincide ni con la falsedad ni con ninguna de las otras propiedades del enunciado que formula el mentiroso, pero que requiere que el mentiroso produzca su enunciado con determinado estado de ánimo, es decir, con la intención de engañar.

Otra cuestión es si alguna característica esencial del *humbug* o de la mentira *no* depende de las intenciones y las creencias de la persona responsable del *humbug* o de la mentira, o si, por el contrario, es posible que un enunciado cualquiera –siempre que el hablante tenga cierta disposición– sea un vehículo del *humbug* o de la mentira. En algunos análisis de la mentira, la mentira no existe a menos que se emita un enunciado falso; en otros, una persona puede estar mintiendo incluso si el enunciado que produce es ver-

2 M. Black, *The prevalence...*, *op. cit.*, p. 143.

dadero, siempre que crea que el enunciado es falso y tenga la intención de engañar al producirlo. ¿Qué se puede decir del *humbug* y del *bullshit*? ¿Se puede considerar *humbug* o *bullshit* cualquier enunciado, siempre que (por así decirlo) quien lo emita tenga la disposición de ánimo que corresponde, o el enunciado también debe tener ciertas características propias?

Que no llega a ser una mentira: debe decirse que *humbug* "no llega a ser mentira", ya que si bien tiene algunas de las características distintivas de las mentiras, carece de otras. Sin embargo, ésta no puede ser toda la idea. Después de todo, el uso del lenguaje, sin excepción, comparte algunas características distintivas de las mentiras, aunque no todas: al menos, en principio, la característica de ser un uso del lenguaje. Sin embargo, con seguridad sería incorrecto describir todo uso del lenguaje como algo que no llega a ser una mentira. La frase de Black evoca la noción de una suerte de continuo, en el que la mentira ocupa determinado segmento, mientras que el *humbug* se ubica exclusivamente en puntos anteriores. ¿Qué continuo podría ser éste, a lo largo del cual se encuentra el *humbug* sólo antes de encontrar la mentira? Tanto la mentira como el *humbug* son modos de tergiversación. Sin embargo, a primera vista no es evidente cómo podría interpretarse que la diferencia entre estas variedades de tergiversación es una diferencia de grado.

En especial mediante una palabra o un hecho pretencioso: aquí se debe prestar atención a dos cuestiones. En primer lugar, Black identifica el *humbug* no sólo como una categoría del lenguaje, sino también como una categoría de la acción; se puede producir mediante palabras o hechos. En segundo lugar, el uso que hace del focalizador *en especial* indica que Black no considera el carácter pretencioso una característica esencial o completamente indispensable del *humbug*. Sin duda, una gran cantidad de *humbug* es pretencioso. Más aun, con respecto al *bullshit*, la frase "*bullshit* pretencioso" se asemeja a un cliché. Sin embargo, me inclino a pensar que cuando el *bullshit* es pretencioso, ello se debe a que el carácter pretencioso es su motivo y no un elemento constitutivo de su esencia. El comportamiento pretencioso de una persona no es, en mi opinión, parte de los requerimientos para considerar que su enunciado sea un ejemplo de *bullshit*. Con seguridad, a menudo ese comporta-

miento explica la producción de ese enunciado. No obstante, hay que suponer que el *bullshit* siempre y necesariamente tiene la pretensión como motivo.

Tergiversación [...] de los pensamientos, sentimientos o actitudes de alguien: esta condición que afirma que el perpetrador del *humbug* se está, en esencia, tergiversando a sí mismo da lugar a cuestiones de importancia central. Para comenzar, cuando una persona tergiversa *algo* deliberadamente, en forma inevitable debe tergiversar su propia disposición de ánimo. Es posible, por supuesto, que una persona sólo tergiverse eso: por ejemplo, fingiendo tener un deseo o un sentimiento que en realidad no tiene. Pero supongamos que una persona, ya sea diciendo una mentira o de otra manera, tergiversa alguna otra cosa. Entonces necesariamente tergiversa, al menos, dos cosas. Tergiversa aquello sobre lo que está hablando, es decir, la situación que es el tópico o el referente de su discurso y, al hacerlo, no puede evitar tergiversar también su propio parecer. De esta manera, alguien que miente acerca de cuánto dinero tiene en el bolsillo da una versión de la cantidad de dinero que tiene en el bolsillo y, a la vez, comunica que cree en esa versión. Si la mentira funciona, entonces su víctima es engañada doblemente, pues tiene una creencia falsa sobre lo que hay en el bolsillo del mentiroso y otra creencia falsa sobre lo que el mentiroso tiene en mente.

Ahora bien, es improbable que Black desee que el referente del *humbug* sea, en todos los casos, la disposición de ánimo del hablante. Después de todo, no hay ninguna razón particular por la cual el *humbug* no podría referirse a otras cosas. Es probable que Black quiera decir que el *humbug* no está destinado, ante todo, a transmitir al oyente una creencia falsa acerca de la situación sobre la que se está hablando, sino que su intención principal es, en cambio, dar a su audiencia una impresión falsa respecto de lo que está sucediendo en la mente del hablante. En la medida en que sea *humbug*, la creación de esta impresión es su objetivo principal y su razón de ser.

Interpretar a Black de este modo sugiere una hipótesis que puede justificar su caracterización del *humbug* como algo "que no llega a ser una mentira". Si le miento a usted acerca de la cantidad de dinero que tengo, no estoy haciendo una afirmación *explícita*

sobre mis creencias. Por tanto, se podría sostener razonablemente que, aunque al decir la mentira en realidad tergiverso lo que tengo en mente, esta tergiversación –a diferencia de mi tergiversación sobre lo que tengo en el bolsillo– no es, en sentido estricto, una mentira. Ello se debe a que no suelto cualquier enunciado acerca de lo que tengo en mente. Tampoco el enunciado que emito, por ejemplo: "Tengo veinte dólares en el bolsillo", entraña un enunciado que me atribuya una creencia. Por otra parte, es incuestionable que al afirmarlo, le brindo a usted un fundamento razonable para emitir ciertos juicios acerca de lo que creo. En particular, le proveo un fundamento razonable para suponer que creo que hay veinte dólares en mi bolsillo. Debido a que esta suposición es, por hipótesis, falsa, al decir la mentira tiendo a engañarlo respecto de lo que tengo en mente aun cuando en realidad no estoy mintiendo acerca de ello. Desde este punto de vista, no parece antinatural ni inapropiado considerar que estoy tergiversando mis propias creencias de una manera tal que "no llega a ser una mentira".

Es fácil pensar en situaciones conocidas que parecen confirmar sin problemas la caracterización que Black hace del *humbug*. Consideremos un orador de un 4 de Julio quien, con grandilocuencia, se refiere a "nuestro bendito y gran país, cuyos Padres Fundadores bajo la guía divina crearon un nuevo comienzo para la humanidad". Sin dudas, esto es *humbug*. Tal como sugiere la caracterización de Black, el orador no está mintiendo. Estaría mintiendo sólo si fuera su intención inducir en su audiencia creencias que él mismo considera falsas, sobre asuntos tales como si nuestro país es grande, si está bendito, si los Fundadores tenían una guía divina y si lo que hicieron fue, de hecho, crear un nuevo comienzo para la humanidad. Sin embargo, al orador no le importa realmente qué piensa su audiencia acerca de los Padres Fundadores o acerca de la función de la deidad en la historia de nuestro país o cosas similares. Al menos, lo que motiva su discurso no es su interés en lo que los demás piensan acerca de estos asuntos.

Está claro que lo que convierte el discurso del 4 de Julio en *humbug* no es fundamentalmente el hecho de que el orador considere que sus afirmaciones son falsas. En cambio, tal como sugiere la caracterización de Black, el orador pretende que estas afirmaciones

transmitan cierta impresión de él mismo. No está tratando de engañar a nadie en cuanto a la historia estadounidense. Lo que le importa es lo que la gente piensa de *él.* Quiere que piensen que es un patriota, alguien que tiene pensamientos y sentimientos profundos acerca de los orígenes y la misión de nuestro país, que aprecia la importancia de la religión, que es sensible a la grandeza de nuestra historia, cuyo orgullo por esa historia se combina con la humildad ante Dios, y así sucesivamente.

La caracterización del *humbug* de Black parece, entonces, ajustarse bastante bien a ciertos paradigmas. No obstante, no creo que capte en forma adecuada o precisa el carácter esencial del *bullshit.* Es correcto decir del *bullshit* –como él dice del *humbug*– a la vez que no llega a ser una mentira, y que aquellos que lo perpetran en cierta manera se tergiversan a sí mismos. Sin embargo, la versión de Black sobre estas dos características es bastante desacertada. A continuación intentaré desarrollar, teniendo en cuenta cierto material biográfico de Ludwig Wittgenstein, una apreciación preliminar, aunque más exacta, de cuáles son las características principales del *bullshit.*

Una vez Wittgenstein dijo que el siguiente fragmento de un verso de Longfellow podía servirle como lema:[3]

> In the elder days of art
> Builders wrought with greatest care
> Each minute and unseen part,
> For the Gods are everywhere.
>
> [En los antiguos tiempos del arte,
> los creadores forjaron con gran cuidado
> cada parte diminuta e invisible
> pues los dioses están en todos lados.]

El sentido de estas líneas es claro. En la Antigüedad, los artesanos no elegían el camino más corto. Trabajaban con esmero, y tenían

3 Citado por Norman Malcolm en su Introducción a R. Rhees (comp.), *Recollections of Wittgenstein*, Oxford, Oxford University Press, 1984, p. XIII [trad. esp.: *Recuerdos de Wittgenstein*, México, Fondo de Cultura Económica, 1989].

cuidado con todos los aspectos de su trabajo. Se tenía en cuenta cada parte del producto, y cada una era diseñada y confeccionada de la manera exacta como debía ser. Estos artesanos no relajaban su atenta autodisciplina ni siquiera respecto de aquellas características de su trabajo que, por lo general, no serían visibles. Aunque nadie lo notara, en el caso de que esos detalles estuvieran descuidados, a los artesanos les pesaría en su conciencia. Por tanto, no se intentaba ocultar nada. O, quizá también se podría decir, no había *bullshit*.

Podría ser adecuado interpretar que los productos confeccionados sin cuidado y de mala calidad son análogos, en cierta forma, al *bullshit*. Pero, ¿de qué manera? La semejanza ¿se debe a que el *bullshit* mismo invariablemente se produce de manera descuidada o permisiva, que nunca se lo confecciona con cuidado, que en su fabricación falta la preocupación atenta y minuciosa por el detalle a la que alude Longfellow? El responsable del *bullshit* ¿es, por naturaleza propia, un patán insensato? Su producto ¿es necesariamente desprolijo o no refinado? La palabra *shit*, con seguridad, lo sugiere. El excremento no es diseñado ni confeccionado; simplemente se lo emite o se lo arroja. Puede tener una forma más o menos coherente, o no, pero, en todo caso, no está en absoluto *forjado*.

La noción del *bullshit* cuidadosamente forjado contiene, entonces, una cierta presión interna. Una atención cuidadosa al detalle requiere disciplina y objetividad. Supone aceptar los estándares y las limitaciones que prohíben la indulgencia del impulso o el capricho. Este desinterés es el que, junto con el *bullshit*, nos resulta inapropiado. Sin embargo, de hecho no es en absoluto imposible. Los ámbitos de la publicidad y las relaciones públicas, y el ámbito de la política –en la actualidad, estrechamente relacionado con ellos– están repletos de ejemplos de *bullshit* tan extremos que se pueden incluir entre los paradigmas más irrefutables y clásicos del concepto. Y en estos ámbitos, se encuentran artesanos exquisitamente sofisticados, que –con la ayuda de técnicas avanzadas y exigentes de investigación de mercado, de encuestas de opinión pública, de pruebas psicológicas y otras– se dedican incansablemente a hacer que cada palabra y cada imagen que producen sea la correcta.

Sin embargo, hay más que añadir al respecto. Por más aplicado y consciente que sea el responsable del *bullshit*, sigue siendo verdad

que también está tratando de actuar con impunidad. Con seguridad, en su trabajo –como en el trabajo de los artesanos descuidados– se encuentra cierto tipo de negligencia que resiste o elude las exigencias de una disciplina desinteresada y austera. El modo de negligencia pertinente no puede equipararse, por cierto, con un simple descuido o con falta de atención al detalle. Intentaré, a su debido tiempo, ubicarlo en una forma más correcta.

Wittgenstein dedicó gran parte de sus energías filosóficas a identificar y combatir lo que consideraba formas insidiosamente perjudiciales de *sinsentido*. Al parecer, también era así en su vida personal. Esto se observa en una anécdota contada por Fania Pascal, quien lo conoció en Cambridge en la década de 1930:

> Me habían operado de las amígdalas y me encontraba en el Evelyn Nursing Home sintiendo pena de mí misma. Wittgenstein vino a visitarme. Yo dije con voz ronca: "Me siento como un perro al que acaban de atropellar". Él se disgustó: "Tú no sabes cómo se siente un perro que acaba de ser atropellado".[4]

Ahora bien, ¿quién sabe qué sucedió en realidad? Parece extraordinario, casi increíble que alguien pueda objetar seriamente la observación que Pascal cuenta que hizo. La caracterización de sus sentimientos –tan inocentemente cercana a la trillada comparación "enfermo como un perro"– no es lo suficientemente provocativa como para dar lugar a una respuesta tan vívida o intensa como el disgusto. Si la comparación de Pascal es ofensiva, entonces, ¿qué usos figurativos o alusivos del lenguaje no lo serían?

Por consiguiente, es posible que, en realidad, la historia no haya sido tal como la cuenta Pascal. Quizá Wittgenstein pretendía hacer un pequeño chiste, y no causó gracia. Sólo fingía que regañaba a Pascal, para divertirse con una pequeña hipérbole; y ella no interpretó bien el tono ni la intención. Ella pensó que él se había disgustado con su observación, cuando de hecho sólo estaba tratando de animarla con una crítica fingida o una chanza exagerada, para jugar. En ese caso, el incidente no es increíble ni extraño después de todo.

4 Fania Pascal, "Wittgenstein: a personal memoir", en R. Rhees, *op. cit.*, pp. 28-29.

Pero si Pascal no pudo darse cuenta de que Wittgenstein sólo estaba bromeando, quizá la posibilidad de que hablara en serio no era, al menos, tan imposible. Ella lo conocía, y sabía qué esperar de él; sabía cómo él la hacía sentir. Probablemente, la forma en que ella interpretó o malinterpretó el comentario de Wittgenstein no discrepaba respecto de la impresión que ella tenía acerca de cómo era él. Sin duda, podemos suponer que aun si su versión del incidente no es estrictamente fiel a la intención de Wittgenstein, es lo suficientemente fiel a la idea que ella tenía de Wittgenstein como para que tuviera sentido para ella. A los fines de este análisis, aceptaré el relato de Pascal, suponiendo que, en relación con el uso de lenguaje alusivo o figurativo, Wittgenstein fue realmente tan absurdo como ella lo hace parecer.

Entonces, ¿qué es lo que el Wittgenstein de su relato considera objetable? Supongamos que él tiene razón acerca de los hechos: es decir, Pascal realmente no sabe cómo se sienten los perros que han sido atropellados. Aun así, cuando ella dice lo que dice, es obvio que no está *mintiendo*. Habría estado mintiendo si, al hacer la afirmación, hubiera sido consciente de que en realidad se sentía bastante bien. Por poco que sepa acerca de la vida de los perros, sin duda Pascal debe de tener en claro que cuando los perros son atropellados no se sienten bien. Por consiguiente, si ella se hubiera estado sintiendo bien, habría mentido al afirmar que se sentía como un perro después de ser atropellado.

El Wittgenstein de Pascal no tiene la intención de acusarla de estar mintiendo, sino de incurrir en una tergiversación de otro tipo. Ella describe su sentimiento como "el sentimiento de un perro que ha sido atropellado". Sin embargo, ella no conoce realmente bien el sentimiento al que esta frase hace referencia. Por supuesto, la frase dista mucho de ser un sinsentido completo para ella; no está diciendo nada incoherente. Lo que dice tiene una connotación inteligible, que por cierto ella comprende. Más aun, ella sabe algo acerca de la cualidad del sentimiento al que se refiere la frase: sabe, al menos, que se trata de un sentimiento indeseable y del que no se disfruta, un sentimiento *malo*. El problema de su enunciado es que pretende transmitir algo más que el simple hecho de sentirse mal. Su descripción del sentimiento es demasiado específica; es excesi-

vamente particular. El suyo no es sólo un malestar cualquiera sino, según su formulación, el tipo de malestar característico que tiene un perro cuando lo atropellan. A juzgar por la respuesta del Wittgenstein del relato de Pascal, para él esto es puro *bullshit.*

Supongamos ahora que Wittgenstein realmente considera un caso de *bullshit* la descripción de Pascal de la forma en que se siente. ¿Por qué le da esa impresión? Creo que se debe a que él percibe que lo que dice Pascal está –en términos generales, por ahora– desvinculado de una preocupación con la verdad. Su afirmación no guarda relación con la tarea de describir la realidad. Ella ni siquiera piensa que sabe, excepto de manera sumamente vaga, cómo se siente un perro que ha sido atropellado. La descripción que hace de su propio sentimiento es, por consiguiente, algo que ella inventa. Es pura invención; o, si alguna otra persona se lo dijo, está repitiéndolo en forma mecánica y sin ninguna consideración por cómo son las cosas en realidad.

Wittgenstein censura a Pascal por esa falta de reflexión. Lo que lo indigna es que a Pascal ni siquiera la preocupa si el enunciado es correcto. Hay muchas probabilidades, por supuesto, de que ella diga lo que dice sólo en un intento algo burdo de hablar con gran colorido o de parecer vivaz y de buen humor; y, sin duda, la reacción de Wittgenstein –según ella la interpreta– es absurdamente intolerante. Sea como fuere, parece claro cuál es esa reacción. Él reacciona como si percibiera que ella está hablando acerca del sentimiento que tiene sin pensar y sin prestar una atención consciente a los hechos pertinentes. Su enunciado no ha sido "forjado con gran cuidado". Lo emite sin preocuparse por tener en cuenta en absoluto la cuestión de su exactitud.

El asunto que le molesta a Wittgenstein no es, evidentemente, que Pascal haya cometido un error en su descripción de cómo se siente. Ni siquiera es que haya cometido un error por descuido. Su negligencia, o su falta de cuidado, no consiste en haber permitido que se deslizara un error en su discurso debido a un lapsus involuntario o momentáneamente negligente en la atención que estaba poniendo para hacer las cosas bien. Más bien, la cuestión es, hasta donde puede ver Wittgenstein, que Pascal ofrece una descripción de cierta situación sin someterse en forma genuina a las restriccio-

nes que impone el esfuerzo tendiente a proporcionar una representación precisa de la realidad. Su falta no reside en no poder hacer los cosas bien, sino en el hecho de que ni siquiera lo intenta.

Esto es importante para Wittgenstein porque, justificadamente o no, toma en serio lo que ella dice, como un enunciado que pretende dar una descripción informativa de la manera en que ella se siente. Él interpreta que ella está involucrada en una actividad para la cual la distinción entre lo que es verdad y lo que es falso es crucial y, sin embargo, que no le interesa si lo que dice es verdadero o falso. En este sentido, el enunciado de Pascal está desvinculado de una preocupación por la verdad: a ella no le interesa el valor de verdad de lo que dice. Por esa razón, no se puede considerar que esté mintiendo, ya que ella no supone que conoce la verdad y, por tanto, no puede estar promulgando en forma deliberada una proposición que ella supone que es falsa. Su enunciado no se basa en la creencia de que es verdad ni, tal como debe ser una mentira, en la creencia de que no es verdad. Esa falta de conexión con una preocupación por la verdad –esta indiferencia respecto de cómo son las cosas en realidad– es lo que yo considero la esencia del *bullshit*.

Ahora analizaré (en forma bastante selectiva) algunas entradas del *Oxford English Dictionary* que permiten elucidar la naturaleza de *bullshit*. El OED define una *bull session** como "una conversación o charla informal, en especial de un grupo de varones". Ahora bien, como definición, parece errónea. En primer lugar, es evidente que el diccionario supone que el uso del término *bull* ["toro"] en *bull session* sirve, fundamentalmente, sólo para indicar el género. Pero incluso si fuera verdad que los participantes de las *bull sessions* son en general o habitualmente varones, la afirmación de que una *bull session* es, en esencia, nada más particular que una charla informal entre varones sería tan desacertada como la afirmación paralela de que una *hen session*** es simplemente una conversación informal entre mujeres. Probablemente sea verdad que las participantes de las *hen sessions* tengan que ser mujeres. No obstante, el término *hen*

* Literalmente, "sesión de toros". [N. de T.]
** Literalmente, "sesión de gallinas". [N. de T.]

session expresa algo más específico con respecto al tipo particular de conversación informal entre mujeres que se desarrolla, por lo general, en las *hen sessions*. El rasgo distintivo del tipo de conversación informal entre varones que constituye una *bull session* es, me parece, algo así: si bien la conversación puede ser intensa y significativa, en cierto sentido no va "en serio".

Los temas característicos de una *bull session* tienen que ver con aspectos de la vida muy personales y cargados de emoción, por ejemplo, la religión, la política o el sexo. Por lo general, las personas son reacias a hablar abiertamente acerca de estos temas si suponen que se las va a tomar demasiado en serio. Lo que suele suceder en una *bull session* es que los participantes prueban diversos pensamientos y actitudes para ver cómo se sienten al oírse decir tales cosas y para descubrir cómo responden los demás, sin que se suponga que esto los compromete con lo que dicen. En una *bull session*, todos comprenden que los enunciados que hace la gente no necesariamente revelan lo que en realidad creen o sienten. La cuestión principal es posibilitar un alto nivel de franqueza y un enfoque experimental o aventurero respecto de los temas que se debaten. Por tanto, se toman las previsiones para gozar de cierta irresponsabilidad, de manera que las personas se vean alentadas a transmitir lo que tienen en mente sin demasiada angustia por tener que atenerse a su palabra.

Dicho de otra manera, todos los que participan en una *bull session* confían en un reconocimiento general de que lo que expresan o dicen no se debe interpretar como lo que sinceramente quieren decir o creen en forma inequívoca que sea verdad. El objetivo de la conversación no es comunicar creencias. Por consiguiente, las suposiciones habituales acerca de la conexión entre lo que la gente dice y lo que cree quedan en suspenso. Los enunciados emitidos en una *bull session* difieren del *bullshit* en que no se pretende que se mantenga esta conexión. Son como el *bullshit* porque, hasta cierto punto, no están condicionados por una preocupación por la verdad. Esta similitud entre las *bull sessions* y el *bullshit* también está sugerida por la expresión *shooting the bull*,* que se refiere al tipo de

* Literalmente, "dispararle al toro"; en sentido figurado, "darle a la lengua". [N. de T.]

conversación que caracteriza a las *bull sessions* y en las que el término *shooting* ("disparar") es muy probablemente una versión lavada de *shitting* ("cagar"). Lo más probable es que la propia expresión *bull session* sea una versión aséptica de *bullshit session.*

Un tema similar se puede discernir en un uso británico de *bull* en el que, según el *OED*, el término se refiere a un "ritual o tareas de rutina innecesarias; disciplina excesiva o pulcritud; burocracia". El diccionario brinda los siguientes ejemplos de este uso:

> "El escuadrón [...] se sintió muy bolche respecto de todas esas porquerías [*bull*] que volaban por la estación" (I. Gleed, *Arise to Conquer* VI. 51, 1942); "Ellos, formando la guardia; nosotros, marchando delante de ellos con la vista hacia la derecha, ese tipo de tonterías [*bull*]" (A. Baron, *Human kind* XXIV. 178, 1953); "El trabajo monótono y las 'sandeces' [*bull*] en la vida de un miembro del Parlamento" (*Economist*, 8 de febrero, 470/471, 1958).

Aquí, el término *bull* se relaciona, evidentemente, con tareas inútiles, en el sentido de que no tienen mucho que ver con la intención primaria o el objetivo que justifica la empresa que las requiere. Se supone que la pulcritud y la burocracia no contribuyen de manera genuina a los objetivos "reales" del personal de las Fuerzas Armadas o los funcionarios del gobierno, aunque sean impuestos por organismos o agentes que dicen estar conscientemente dedicados a la búsqueda de esos objetivos. Por tanto, "el ritual o las tareas de rutina innecesarias" que constituyen el *bull* no se relacionan con los motivos que legitiman las actividades en las que se inmiscuyen, de la misma manera que las cosas que las personas dicen en las *bull sessions* no se relacionan con sus creencias establecidas y que el *bullshit* no se relaciona con una preocupación por la verdad.

El término *bull* también se emplea, en un uso algo más generalizado y conocido, como un equivalente algo menos grosero de *bullshit*. En una acepción de *bull* referida a ese uso, el *OED* sugiere la siguiente explicación como definitoria: "charla o escrito trivial, insincero o falso; tonterías". Ahora bien, no parece un rasgo distintivo de *bull* el que tenga que ser deficiente en significado o necesariamente poco importante; por tanto, *tonterías* y *trivial*, aun de-

jando a un lado su vaguedad, parecen estar en la senda equivocada. El matiz de *insincero o falso* es mejor, aunque debe precisarse.[5] La entrada también brinda las dos definiciones siguientes:

> 1914 *Dialect Notes* IV. 162 *Bull*, palabras que no vienen al caso; palabrería.
> 1932 *Times Lit. Supp.* 8 de diciembre. 933/3 *Bull* es el término en argot empleado para una combinación de *bluff*,* bravuconadas, palabrería y lo que solíamos llamar en el Ejército "tomarle el pelo a la tropa".

La expresión "no viene al caso" es adecuada, pero demasiado vaga y con un alcance demasiado amplio. Abarca las digresiones y las irrelevancias inocentes, que no siempre son casos de *bull*; más aun, decir que el *bull* no viene al caso deja su propósito sin precisar. La referencia a la palabrería en ambas definiciones resulta más útil.

Cuando describimos la charla como palabrería, lo que queremos decir es que lo que sale de la boca del hablante es meramente eso. Son sólo palabras. Su discurso está vacío, no tiene ni sustancia ni contenido. Su uso del lenguaje, por tanto, no contribuye al objetivo al que pretende servir. No se comunica más información que la que el hablante habría transmitido con una mera exhalación. Hay similitudes entre la palabrería y el excremento, por cierto, lo que hace que la *palabrería* aparezca como un equivalente especialmente adecuado de *bullshit*. Así como la palabrería es un discurso que ha sido vaciado de todo contenido informativo, el excremento es una materia que ha perdido todos los elementos nutritivos. El excremento puede considerarse el cadáver del alimento, lo que queda cuando los elementos vitales de la comida se han agotado. En este aspecto, el excremento es una representación de la muerte que nosotros mismos producimos y que, de hecho, no podemos evitar producir en el proceso mismo de mantener nuestras vidas. Quizá porque hace que la muerte sea tan íntima es

5 Se observará que la inclusión de la falta de sinceridad entre sus condiciones esenciales implicaría que el *bull* no se puede producir de manera inadvertida, puesto que no parece posible ser inadvertidamente insincero.

* Mantenemos la palabra en inglés, de uso habitual también en español. [N. de T.]

que el excremento nos resulta tan repulsivo. De todas maneras, no sirve para el sustento, no más que la palabrería puede servir para la comunicación.

Ahora consideremos estas líneas del Canto LXXIV de Pound, que el *OED* cita en su entrada de *bullshit* como verbo:

> Hey Snag wots in the bibl'?
> Wot are the books ov the bible?
> Name 'em, don't bullshit ME.[6]
>
> [Oye, Snag, ¿qué hay en la biblia?
> ¿Cuáles son los libros de la biblia?
> Nómbramelos y no me jodas a MÍ.]

Esto es un llamado a los hechos. Evidentemente, se considera que el interlocutor ha afirmado, de cierta manera, conocer la Biblia o ha afirmado que ésta le importa. El hablante sospecha que sólo se trata de palabras vacías, y exige que la afirmación se respalde con hechos. No aceptará un mero relato, insiste en ver la cosa en sí misma. En otras palabras, está denunciando su juego (*calling the bluff*). La conexión entre *bullshit* y *bluff* se afirma de manera explícita en la definición con la que se asocian las líneas de Pound:

6 A continuación incluyo parte del contexto en que estas líneas ocurren: "Les Albigeois, a problem of history, / and the fleet at Salamis made with money lent by the state to the shipwrights / Tempus tacendi, tempus loquendi. / Never inside the country to raise the standard of living / but always abroad to increase the profits of usurers, / dixit Lenin, / and gun sales lead to more gun sales / they do not clutter the market for gunnery / there is no saturation / Pisa, in the 23rd year of the effort in sight of the tower / and Till was hung yestarday / for murder and rape with trimmings plus Cholkis / plus mythology, thought he was Zeus ram or another one / Hey Snag wots in the bibl'? / wot are the books ov the bible? Name 'em, don't bullshit ME". [Los albigenses, un problema de la historia / y la flota de Salamina construida con dinero prestado por el Estado a los armadores / Tempus tacendi, tempus loquendi. / Jamás dentro del país para mejorar el nivel de vida, / sino siempre en el extranjero para acrecer las ganancias de los usureros, / dixit Lenin, / y el vender armas lleva a vender más armas / no se satura el mercado de cañones / no hay saturación / Pisa en el año 23º del esfuerzo a la vista de la torre / y ayer colgaron a Till / por asesinato y violación con accesorios más Cólquide / más mitología, creía que era el Zeus carnero u otro / Oye, Snag, ¿qué hay en la biblia? / ¿Cuáles son los libros de la biblia? / Nómbramelos y no me jodas a MÍ.]

> Como verbo transitivo e intransitivo, decir tonterías; también [...], *salir del apuro* (*to bluff one's way through something*) diciendo tonterías.

Parecería que la producción de *bullshit* implica un tipo de *bluff*. Sin duda se parece más a producir un *bluff* que a mentir. Pero, ¿qué consecuencias tiene, en cuanto a su naturaleza, el hecho de parecerse más al primer caso que al segundo? ¿Cuál es la diferencia relevante entre un *bluff* y una mentira?

Mentir y producir *bluffs* son dos modos de tergiversación o de engaño. Ahora bien, el concepto particularmente central para la naturaleza distintiva de una mentira es el de la falsedad; un mentiroso es, en esencia, alguien que promulga una falsedad en forma deliberada. De manera característica, el producir un *bluff* también está destinado a transmitir algo falso. A diferencia de la pura mentira, sin embargo, no es en especial una cuestión de falsedad, sino de falsificación. Esto es lo que explica su proximidad respecto del *bullshit*. Ello se debe a que la esencia del *bullshit* no es que sea *falso* (*false*), sino *falsificado* (*phony*). A fin de apreciar esta distinción, debemos reconocer que algo falso o falsificado no necesita ser, en ningún aspecto (aparte de la autenticidad misma), inferior a la cosa real. Lo que no es genuino no necesariamente tiene que ser defectuoso en ningún otro sentido. Puede ser, después de todo, una copia exacta. Lo incorrecto de una imitación no es su aspecto o su calidad, sino el modo en que fue hecha. Esto apunta a un aspecto similar y básico de la naturaleza esencial del *bullshit:* aunque se lo produce sin preocupación por la verdad, no necesariamente es falso. El productor de *bullshit* está falsificando las cosas. Sin embargo, esto no significa que necesariamente las haga mal.

En la novela de Eric Ambler, *Dirty Story*, un personaje llamado Arthur Abdel Simpson recuerda el consejo que su padre le había dado cuando era un niño:

> Aunque sólo tenía siete años de edad cuando mataron a mi padre, aún lo recuerdo muy bien, así como también algunas de las cosas que solía decir [...]. Una de las primeras cosas que me en-

señó fue: *Nunca digas una mentira cuando puedas salir del apuro mediante el bullshit.*[7]

Esto supone no sólo que hay una importante diferencia entre mentir y producir *bullshit*, sino que lo último es preferible a lo primero. Ahora bien, seguramente Simpson padre no consideraba que producir *bullshit* era moralmente superior a mentir. Tampoco es probable que considerara que las mentiras fueran invariablemente menos efectivas que el *bullshit* a fin de lograr los objetivos para los cuales se podrían usar unas u otro. Después de todo, una mentira inteligentemente elaborada puede cumplir su función con total éxito. Puede que Simpson pensara que es más fácil salir del apuro mediante el *bullshit* que mintiendo. O quizá quería decir que, aunque el riesgo de ser atrapado es aproximadamente el mismo en ambos casos, las consecuencias de ser atrapado por lo general son menos severas para el responsable del *bullshit* que para el mentiroso. De hecho, las personas tendemos a tolerar más el *bullshit* que las mentiras, quizá porque nos sentimos menos inclinadas a tomar el primero como una afrenta personal. Puede ser que busquemos distanciarnos del *bullshit*, pero es más probable que nos alejemos de él encogiendo los hombros con impaciencia o irritación que con el sentimiento de violación o de furia que suelen inspirar las mentiras. El problema de comprender por qué nuestra actitud hacia el *bullshit* es, por lo general, más benigna que nuestra actitud hacia la mentira es importante, y lo dejaré como ejercicio para el lector.

No obstante, la comparación pertinente no es entre decir una mentira y producir alguna instancia particular de *bullshit*. Simpson padre identifica la alternativa a la mentira con la frase "salir del apuro mediante el *bullshit*". Esto supone no sólo producir una instancia de *bullshit*; supone un *programa* de producción de *bullshit* en la medida requerida por las circunstancias. Ésta puede ser una

7 E. Ambler, *Dirty story*, 1967, I. III. 25. La cita se encuentra en la misma entrada del OED que incluía el pasaje de Pound. A mi entender, la proximidad de la relación entre *bullshit* y *bluff* es resonante en el paralelismo de las expresiones idiomáticas: "bullshit your way through" [salir del apuro mediante el *bullshit*] y "bluff your way through" [salir del apuro mediante el *bluff*].

clave de su preferencia. Decir una mentira es un acto con un objetivo preciso. Está destinado a insertar una falsedad en particular en un punto específico dentro de un conjunto o sistema de creencias a fin de evitar las consecuencias que tendría el hecho de ocupar ese punto con la verdad. Esto requiere un grado de destreza, en el cual quien dice la mentira se somete a las restricciones objetivas impuestas por lo que él piensa que es la verdad. El mentiroso está inevitablemente interesado en los valores de verdad. A fin de inventar una mentira cualquiera, debe pensar que conoce la verdad. Y para inventar una mentira efectiva, debe diseñar su falsedad bajo la guía de esa verdad.

Por otra parte, una persona que pretende salir del apuro mediante el *bullshit* tiene una libertad mucho mayor. Su enfoque es panorámico y no particular. No se limita a insertar determinada falsedad en un punto específico y, por tanto, no está restringida por las verdades que rodean o cruzan ese punto. Asimismo, está preparada para falsificar el contexto en la medida en que sea necesario. Esta libertad respecto de las restricciones a las cuales se debe someter el mentiroso no significa necesariamente, por supuesto, que su tarea sea más fácil que la tarea del mentiroso. Sin embargo, el modo de creatividad sobre el que ella se apoya es menos analítico y deliberativo que el que se moviliza al mentir. Es más expansivo e independiente, y presenta oportunidades más amplias para la improvisación, el color y el juego imaginativo. No se trata tanto de un asunto de artesanía como de arte. De allí, la conocida noción del "artista del *bullshit*". Mi conjetura es que la recomendación ofrecida por el padre de Arthur Simpson refleja el hecho de que estaba impulsado con más fuerza a este modo de creatividad, sin importar su mérito o su efectividad relativos, de lo que estaba a las exigencias más austeras y rigurosas de la mentira.

Lo que el *bullshit* esencialmente tergiversa no es la situación a la que se refiere ni las creencias del hablante respecto de esa situación. Las mentiras son las que las tergiversan, puesto que son falsas. Debido a que no es necesario que el *bullshit* sea falso, difiere de las mentiras en su intención de tergiversar. Puede que el productor de *bullshit* no nos engañe o ni siquiera intente hacerlo, ni respecto de los hechos ni de lo que cree que son los hechos. Sin embargo,

necesariamente intenta engañarnos acerca de su empresa. Su única característica indispensablemente distintiva es que, en cierta manera, tergiversa lo que está haciendo.

Éste es el quid de la distinción entre él y el mentiroso. Tanto él como el mentiroso se representan a sí mismos, falsamente, como personas que intentan comunicar la verdad. El éxito de cada uno depende de que puedan engañarnos sobre ese punto. Pero el hecho que el mentiroso esconde acerca de sí mismo es que está intentando alejarnos de una percepción correcta de la realidad; nosotros no deberíamos saber que él quiere que creamos algo que él supone falso. En cambio, el hecho que el responsable del *bullshit* esconde acerca de sí mismo es que el valor de verdad de sus afirmaciones no es su interés principal; lo que no debemos entender es que su intención no es ni informar la verdad ni ocultarla. Esto no significa que su discurso sea anárquicamente impulsivo, sino que el motivo que lo guía y lo controla no se preocupa por cómo son en realidad las cosas acerca de las que habla.

Es imposible que alguien mienta, a menos que crea conocer la verdad. Producir *bullshit* no requiere una convicción de ese tipo. Una persona que miente está, por ese hecho, respondiendo a la verdad y, en esa medida, la respeta. Cuando habla un hombre honesto, dice sólo lo que cree que es verdad; por tanto, cuando habla el mentiroso es indispensable que considere que sus enunciados son falsos. Sin embargo, para el productor de *bullshit* ninguna de estas opciones vale: no está del lado de la verdad ni del lado de lo falso. Su mirada no está para nada dirigida a los hechos, como sí lo están la mirada de un hombre honesto y la de un mentiroso, excepto en la medida en que esos hechos sean pertinentes para su objetivo de salirse con la suya. No le importa si las cosas que dice describen la realidad correctamente. Sólo las elige o las inventa a fin de que le sirvan para satisfacer su objetivo.

En su ensayo "La mentira", San Agustín distingue ocho tipos de mentiras, que clasifica según la intención o la justificación características con las que se dice una mentira. Las mentiras de siete de estos tipos se dicen sólo porque se supone que son medios indispensables para cierto fin distinto de la mera creación de creencias falsas. En otras palabras, no es su falsedad como tal

lo que atrae al que las dice. Debido a que se dicen sólo a causa de su supuesto carácter indispensable para cierto fin distinto del engaño mismo, San Agustín considera que se dicen con renuencia: lo que la persona realmente quiere no es decir la mentira, sino lograr su objetivo. Por consiguiente, en su opinión, no son mentiras reales y quienes las dicen no son, en el sentido más estricto, mentirosos. Sólo la categoría restante contiene lo que él identifica como "la mentira que se dice sólo por el placer de mentir y engañar, es decir, la auténtica mentira".[8] Las mentiras de esta categoría no se dicen como medio para otro fin que no sea propagar la falsedad. Se dicen simplemente por sí mismas, es decir, por puro amor al engaño:

> Existe una distinción entre una persona que dice una mentira y un mentiroso. El primero es alguien que dice una mentira de mala gana, mientras que al mentiroso le agrada mentir y pasa su tiempo disfrutando de la mentira [...]. El último se complace en mentir, regocijándose en la falsedad misma.[9]

Lo que San Agustín denomina *mentirosos* y *auntênticas mentiras* son a la vez poco comunes y extraordinarios. Todos mienten de tanto en tanto, pero hay muy pocas personas a las que se les ocurriría con frecuencia (o incluso alguna vez) mentir exclusivamente por amor a la falsedad o al engaño.

Para la mayoría de las personas, el hecho de que una afirmación sea falsa constituye, en sí misma, una razón –por más débil y fácilmente invalidada que resulte– para no hacer esa afirmación. Por el contrario, para el mentiroso puro de San Agustín es una razón a favor de hacerla. Para el productor de *bullshit* no es, en sí misma, una razón ni a favor ni en contra. Tanto al mentir como al decir la verdad, las personas son guiadas por sus creencias res-

8 "Lying" [título original: *De Mendacio*], en *Treatises on various subjects*, en R. J. Deferrari (ed.), *Fathers of the Church*, Nueva York, 1952, vol. 16, p. 109. San Agustín sostiene que decir una mentira de este tipo es un pecado menos grave que decir mentiras de tres de sus categorías y un pecado más grave que decir mentiras de las otras cuatro categorías.

9 *Ibid.*, p. 79.

pecto de cómo son las cosas. Éstas las guían ya en su intento de describir el mundo correctamente, ya en el de describirlo de manera engañosa. Por esta razón, decir mentiras no tiende a incapacitar a una persona para decir la verdad de la misma manera en que lo hace el producir *bullshit*. Una complacencia excesiva en esta última actividad –que supone hacer afirmaciones sin estar atento a nada más que lo que a uno le conviene decir– puede hacer que el hábito normal de una persona de prestar atención a cómo son las cosas se atenúe o se pierda. Alguien que miente y alguien que dice la verdad juegan en equipos contrarios, por así decirlo, en el mismo partido. Cada uno responde a los hechos según los entiende, aunque la respuesta de uno está guiada por la autoridad de la verdad, mientras que la respuesta del otro desafía esa autoridad y se niega a satisfacer sus demandas. El productor de *bullshit* hace caso omiso de estas demandas. No rechaza la autoridad de la verdad, como lo hace el mentiroso, ni se opone a ella. No le presta atención en absoluto. En virtud de ello, el *bullshit* es un enemigo de la verdad más poderoso que las mentiras.

Una persona interesada en informar o en esconder los hechos supone que, en realidad, hay hechos que de alguna manera son a la vez determinados y cognoscibles. Su interés en decir la verdad o en mentir presupone que existe una diferencia entre hacer las cosas mal y hacerlas bien, y que es posible, al menos en ocasiones, establecer la diferencia. Alguien que deja de creer en la posibilidad de identificar ciertos enunciados como verdaderos y otros como falsos sólo puede tener dos alternativas. La primera es desistir tanto de los esfuerzos por decir la verdad como de los esfuerzos por engañar. Esto significaría abstenerse de hacer cualquier afirmación sobre los hechos. La segunda alternativa es continuar haciendo afirmaciones que pretenden describir la manera en que son las cosas, pero que no pueden ser nada más que *bullshit*.

¿Por qué hay tanto *bullshit*? Por supuesto que es imposible estar seguros de que hay relativamente más *bullshit* en la actualidad que en otras épocas. Hay más comunicación de todo tipo en nuestro tiempo de lo que nunca antes ha habido, pero es posible que la proporción de *bullshit* no haya aumentado. Sin suponer que la incidencia del *bullshit* sea ahora, en efecto, mayor, mencionaré algu-

nas consideraciones que ayudan a explicar el hecho de que en la actualidad sea tanta.

El *bullshit* es inevitable cuando las circunstancias requieren que alguien hable sin saber de qué está hablando. Así, la producción de *bullshit* es estimulada cada vez que las obligaciones o las oportunidades que tiene una persona para hablar acerca de algún tema son más amplias que su conocimiento de los hechos relevantes para ese tema. Esta discrepancia es común en la vida pública, donde con frecuencia las personas se ven impelidas –ya sea por propensión propia, ya por las demandas de los otros– a hablar extensivamente acerca de asuntos que, hasta cierto punto, ignoran. Otros casos, estrechamente relacionados con los anteriores, surgen de la convicción generalizada de que los ciudadanos de una democracia tienen la responsabilidad de poseer una opinión acerca de todo o, al menos, de todo lo vinculado con la conducción de los asuntos del país. La carencia de una conexión significativa entre las opiniones de una persona y su percepción de la realidad será incluso más grave, no hace falta decirlo, para alguien que cree que es su responsabilidad, como agente moral consciente, evaluar los acontecimientos y las condiciones de todas las regiones del mundo.

La proliferación contemporánea del *bullshit* también tiene fuentes más profundas en diversas formas de escepticismo que niegan que podamos tener un acceso confiable a una realidad objetiva y que, por tanto, rechazan la posibilidad de saber cómo son las cosas en realidad. Estas doctrinas *antirrealistas* debilitan la confianza en el valor de los esfuerzos desinteresados por determinar qué es verdadero y qué es falso e, incluso, en la inteligibilidad de la noción de la investigación objetiva. Una respuesta a esta pérdida de confianza ha sido un alejamiento de la disciplina, requerido por la dedicación al ideal de la *corrección*, hacia un tipo de disciplina bastante diferente, impuesto por la búsqueda de un ideal alternativo de la *sinceridad.* En vez de intentar, en primer lugar, llegar a representaciones precisas de un mundo común, el individuo se vuelca a tratar de brindar representaciones honestas de sí mismo. Convencido de que la realidad no tiene naturaleza inherente, que podría esperar identificar como la verdad de las cosas, se dedica a ser fiel a su propia naturaleza. Parecería decidir que, debido a que no

tiene sentido tratar de ser fiel a los hechos, debe intentar, por tanto, ser fiel a sí mismo.

No obstante, es absurdo imaginar que nosotros mismos somos determinados y, por tanto, susceptibles de descripciones correctas e incorrectas, mientras que suponemos que el hecho de atribuirle la determinación a cualquier otra cosa ha demostrado ser un error. En nuestra condición de seres conscientes, existimos sólo en respuesta a otras cosas, y no podemos conocernos en absoluto sin conocerlas a ellas. Aun más, nada hay en la teoría y, por cierto, tampoco en la experiencia, que respalde el juicio extraordinario de que es la verdad acerca de ella misma la que la persona puede conocer más fácilmente. Los hechos acerca de nosotros mismos no son particularmente sólidos ni resistentes a la disolución escéptica. Nuestra naturaleza es, en efecto, elusivamente insustancial; en forma notoria, menos estable y menos inherente que la naturaleza de otras cosas. Y, en la medida en que éste sea el caso, la sinceridad misma es *bullshit*.

11
La igualdad como ideal moral

Primer Hombre: "¿Cómo están tus hijos?"
Segundo Hombre: "¿Comparados con qué?"

I

El igualitarismo económico es, tal como lo interpretaré, la doctrina que establece que es deseable que todos tengan la misma cantidad de ingresos y de riquezas (en una palabra, *dinero*).[1] Casi nadie negaría que hay situaciones en las que tiene sentido tolerar desviaciones de esta norma. Después de todo, de más está decir que impedir o corregir dichas desviaciones puede acarrear costos que –ya se midan en términos económicos, ya en términos de consideraciones no económicas– son inaceptables según cualquier medida razonable. No obstante, muchas personas creen que la igualdad económica tiene un considerable valor moral en sí misma. Por esta razón, muchas veces insisten en que los esfuerzos para abordar el ideal igualitario deberían recibir una significativa prioridad, con la

1 Esta versión del igualitarismo económico (en una palabra, simplemente *igualitarismo*) también podría formularse como la doctrina de que no debería haber desigualdades en la *distribución* del dinero. Las dos formulaciones no son equivalentes sin ambigüedad, porque el término *distribución* es equívoco. Puede referirse al patrón de la posesión o a una actividad de reparto, y existen diferencias significativas en los criterios para evaluar la distribución en los dos sentidos. Por tanto, es muy posible sostener con coherencia a la vez que es aceptable que las personas tengan cantidades desiguales de dinero y que es objetable repartir el dinero en forma desigual.

debida consideración hacia los efectos que dichos esfuerzos podrían tener de obstruir el logro de otros bienes o de conducir a él.[2]

En mi opinión, esto es un error. La igualdad económica no tiene, como tal, una importancia moral específica. Con respecto a la distribución de los bienes económicos, lo que *es* importante desde el punto de vista de la moralidad no es que todos deban tener *lo mismo*, sino que cada uno tenga *lo suficiente*. Si todos tuvieran lo suficiente, no tendría consecuencias morales el hecho de que algunos tuvieran más que otros. Me referiré a esta alternativa al igualitarismo –es decir, que lo que es moralmente importante con respecto al dinero es que todos tengan lo suficiente– como *la doctrina de la suficiencia*.[3]

Por supuesto, el hecho de que la igualdad económica no sea, por derecho propio, un ideal social moralmente apremiante no es en absoluto una razón para considerar que no sea deseable. Mi afirmación de que la igualdad en sí misma carece de importancia moral no implica que deba evitarse. Sin duda, pueden existir buenas razones para que los gobiernos o los individuos aborden los problemas de la distribución económica de acuerdo con un es-

2 Así, Thomas Nagel escribe: "La defensa de la igualdad económica sobre la base de que es necesaria para proteger la igualdad política, legal y social [no es] una defensa de la igualdad *per se*: la igualdad en la posesión de beneficios en general. Sin embargo, esta última es otra idea moral de gran importancia. Su validez proporcionaría una razón independiente para favorecer la igualdad económica como un bien por derecho propio" ("Equality", en T. Nagel, *Mortal questions*, Cambridge, Cambridge University Press, 1979, p. 107 [trad. esp.: *La muerte en cuestión. Ensayos sobre la vida humana*, México, Fondo de Cultura Económica, 1981]).

3 Concentro mi atención en el estándar de igualdad en la distribución del dinero principalmente a fin de facilitar mi análisis del estándar de suficiencia. Por supuesto, muchos igualitaristas consideran que la igualdad económica es moralmente menos importante que la igualdad en otras cuestiones: por ejemplo, el bienestar, la oportunidad, el respeto, la satisfacción de las necesidades. De hecho, parte de lo que tengo para decir acerca del igualitarismo económico y la suficiencia se aplica también a estos otros beneficios. Sin embargo, no trataré de definir en este ensayo el alcance de su aplicabilidad, como tampoco trataré de expresar mis opiniones con respecto a otra crítica reciente del igualitarismo (por ejemplo, Larry S. Temkin, "Inequality", en *Philosophy and public affairs*, vol. 15, Nº 2, primavera de 1986, pp. 99-121; Robert E. Goodin, "Epiphenomenal egalitarianism", en *Social Research*, vol. 52, Nº 1, primavera de 1985, pp. 99-117).

tándar igualitario y se preocupen más por tratar de aumentar el grado de igualdad de las personas desde un punto de vista económico que intentar regular en forma directa el grado en que la cantidad de dinero que tiene la gente es suficiente. Aunque la igualdad no sea, como tal, moralmente importante, puede ser indispensable que se asuma un compromiso con una política social igualitaria para promover el disfrute de bienes significativos además de la igualdad o para evitar su deterioro. Más aun, podría resultar que el enfoque más viable para el logro de la suficiencia fuera la búsqueda de la igualdad.

Sin embargo, a pesar del hecho de que una distribución igualitaria no necesariamente sería objetable, el error de creer que existen poderosas razones morales para preocuparse por la igualdad dista mucho de ser inocuo. De hecho, esta creencia tiende a causar un daño importante. Muchas veces se esgrime como objeción al igualitarismo el argumento de que existe un peligroso conflicto entre la igualdad y la libertad: si las personas son abandonadas a su libre albedrío, inevitablemente surgen desigualdades de ingresos y de riqueza y, por tanto, una distribución igualitaria del dinero puede lograrse y mantenerse sólo a costa de la represión. Sea cual fuere el mérito de este argumento con respecto a la relación entre la igualdad y la libertad, el igualitarismo económico engendra otro conflicto, de una importancia moral más fundamental.

En la medida en que las personas se preocupen por la igualdad *per se*, su disposición a estar satisfechas con cualquier nivel particular de ingresos o de riqueza no está guiada por sus propios intereses y necesidades, sino sólo por la magnitud de los beneficios económicos que están a disposición de los demás. De esta manera, el igualitarismo distrae a las personas de medir los requisitos a los que dan origen sus naturalezas individuales y sus circunstancias personales. En cambio, las alienta a insistir en un nivel de apoyo económico determinado por un cálculo en el que las características particulares de sus propias vidas son irrelevantes. Después de todo, la magnitud de los bienes económicos de los demás no tiene mucho que ver con la clase de persona que se es. El interés por la igualdad económica, interpretada como deseable en sí misma, tiende a distraer la atención de una persona de tratar de descubrir

–de la experiencia que la persona tiene de sí misma y de su vida– lo que le importa y lo que en realidad lo satisfará, aunque ésta sea la tarea más básica y decisiva de la que depende una selección inteligente de las metas económicas. En otras palabras, exagerar la importancia moral de la igualdad económica es perjudicial, porque es enajenante.[4]

Ciertamente, las circunstancias de los demás pueden revelar posibilidades interesantes y proporcionar información para hacer juicios útiles con respecto a lo que es normal o típico. Esto le puede resultar útil a alguien que trata de alcanzar una apreciación confiada y realista de qué buscar para sí. Más aun, las situaciones de otras personas pueden ser pertinentes, no sólo de modos sugestivos y preliminares como éstos, para alguien que intenta decidir qué exigencias económicas es razonable o importante plantear. La cantidad de dinero que necesita puede depender, de una manera más directa, de la cantidad que tienen los demás. El dinero puede traer poder, prestigio u otras ventajas competitivas. Una persona preocupada con cuestiones de este tipo no puede determinar en forma inteligente cuánto dinero sería suficiente, excepto sobre la base de un cálculo estimativo de los recursos que están disponibles para aquellos con quienes, tal vez, tenga que competir. Sin embargo, lo que es importante desde este punto de vista no es la comparación de los niveles de prosperidad como tales. La medición de la desigualdad es importante sólo en la medida en que se relaciona, en forma contingente, con otros intereses.

El error de creer que la igualdad económica es importante en sí misma lleva a las personas a separar el problema de formular sus ambiciones económicas del problema de comprender qué es lo más fundamental y significativo para ellas. Influye sobre ellas para

4 Podría argumentarse (como me han sugerido algunos de los editores de *Ethics*) que buscar la igualdad como ideal social importante no sería tan enajenante como buscarla como una meta personal. Sin duda, es posible que los individuos dedicados a la primera búsqueda no estuvieran ni tan inmediata ni tan intesamente preocupados por sus propias circunstancias económicas que aquéllos dedicados a la segunda. Sin embargo, apenas considerarían que el logro de la igualdad económica es importante para la sociedad a menos que tuvieran la convicción falsa y enajenante de que sería importante que los individuos gozaran de igualdad económica.

hacerlas tomarse demasiado en serio, como si se tratara de algo de gran importancia moral, una cuestión intrínsecamente insignificante y alejada del punto central: a saber, la comparación de su situación económica con la de los demás. De esta manera, la doctrina de la igualdad contribuye a la desorientación moral y a la superficialidad de nuestra época.

La preponderancia del pensamiento igualitarista es perjudicial también en otro aspecto. No sólo tiende a distraer la atención de las consideraciones que tienen mayor importancia moral que la igualdad. También distrae la atención respecto de los problemas filosóficos difíciles, aunque fundamentales, de comprender cuáles son estas consideraciones y de elaborar, en forma detallada, exhaustiva y perspicua, un aparato conceptual que facilitaría su exploración. Calcular el tamaño de las partes iguales es, simplemente, mucho más fácil que determinar cuánto necesita una persona para tener lo suficiente. Además, el concepto mismo de recibir partes iguales es, en sí, mucho más evidente y accesible que el concepto de tener lo suficiente. De más está decir que no es para nada obvio precisamente qué significa la doctrina de la suficiencia y qué aplicaciones conlleva. Ésta no es, empero, una buena razón para descuidar la doctrina o para adoptar, en cambio, una doctrina incorrecta. Entre mis principales propósitos en este ensayo está el de sugerir la importancia de una investigación sistemática de los temas analíticos y teóricos planteados por el concepto de tener lo suficiente, cuya importancia ha ocultado el igualitarismo.[5]

II

Existen varias maneras de intentar establecer la tesis de que la igualdad económica es importante. A veces se dice con insistencia que la preponderancia de las relaciones fraternales entre los miembros de una sociedad es una meta deseable y que la igualdad es in-

5 Me referiré a algunos de estos temas en la sección VII, más adelante.

dispensable para alcanzarla.[6] O quizá puede sostenerse que las desigualdades en la distribución de los beneficios económicos deben evitarse porque invariablemente conducen a discrepancias no deseadas de otros tipos; por ejemplo, en la condición social, en la influencia política o en la capacidad de la gente de hacer un uso efectivo de sus diversas oportunidades y derechos. En ambos argumentos, se adhiere a la igualdad económica por su supuesta importancia para crear o preservar ciertas condiciones no económicas. Estas consideraciones pueden proporcionar razones convincentes para recomendar la igualdad como un bien social deseable, o incluso para preferir el igualitarismo como política en lugar de las alternativas a aquélla. Sin embargo, ambos argumentos interpretan que la igualdad es valiosa, en forma derivada, en virtud de sus relaciones contingentes con otras cosas. En ninguno de los dos argumentos se le atribuye a la igualdad ningún valor moral inequívocamente intrínseco.

Un argumento de un tipo bastante diferente para la igualdad económica, que se acerca más a interpretar que el valor de la igualdad es independiente de las contingencias, se basa en el principio de la utilidad marginal decreciente. Según este argumento, la igualdad es deseable porque una distribución igualitaria de los bienes económicos maximiza su utilidad agregada.[7] El argumento presupone

6 En la Biblioteca Sterling Memorial de la Universidad de Yale (8,5 millones de volúmenes), hay 1.159 entradas en el catálogo de tarjetas bajo el título *libertad* y 326 bajo *igualdad*. Bajo *fraternidad* no hay ninguno. Ello se debe a que el catálogo se refiere al ideal social en cuestión como *hermandad*. ¡Bajo ese encabezamiento hay *cuatro* entradas! ¿Por qué la fraternidad (o la hermandad) tienen tanta menos importancia que la libertad y la igualdad? Quizá la explicación sea que, en virtud de nuestro compromiso fundamental con el individualismo, los ideales políticos hacia los que nos sentimos más profunda y activamente atraídos tienen que ver con lo que suponemos que son los derechos de los individuos; y nadie exige un derecho a la fraternidad. También es posible que la libertad y la igualdad reciban más atención en ciertos ámbitos porque, a diferencia de la fraternidad, se considera que son susceptibles de un tratamiento más o menos formal. En cualquier caso, el hecho es que ha habido muy poca investigación seria sobre lo que es la fraternidad, lo que ella implica, o por qué debería ser considerada algo especialmente deseable.

7 T. Nagel, en *Mortal questions*, adhiere a este argumento afirmando que establece la importancia moral de la igualdad económica. Pueden encontrarse otras formulaciones y análisis del argumento en Kenneth Arrow, "A utilitarian approach

lo siguiente: (a) para cada individuo, la utilidad del dinero decrece invariablemente en el margen, y (b) con respecto al dinero, o con respecto a las cosas que el dinero puede comprar, las funciones de utilidad de todos los individuos son las mismas.[8] En otras palabras, la utilidad proporcionada por una *enésima* unidad de dinero –o derivada de ella– es la misma para todos, y es menos que la utilidad que (n - 1) unidades de dinero tiene para cualquiera. A menos que (b) fuera verdad, un hombre rico podría obtener una mayor utilidad que un hombre pobre de una unidad de dinero adicional. En ese caso, una distribución igualitaria de los bienes económicos no maximizaría la utilidad agregada aunque (a) fuera verdad. Sin embargo, dados (a) y (b), se sigue que un dólar marginal siempre trae menos utilidad a una persona rica que a otra que es menos rica. Y esto implica que la utilidad total debe aumentar cuando la desigualdad se reduce al entregarle una unidad de dinero a alguien más pobre que la persona que se lo entrega.

De hecho, sin embargo, tanto (a) como (b) son falsos. Supongamos que se reconoce, a los fines del argumento, que la maximización de la utilidad agregada es, por derecho propio, una meta social moralmente importante. Aun así, no puede inferirse en forma legítima que una distribución igualitaria del dinero deba, por tanto, tener una importancia moral similar. Ello se debe a que, en

---

to the concept of equality in public expenditures", en *Quarterly Journal of Economics*, Nº 85, 1971; Walter Blum y Harry Kalven, *The uneasy case for progressive taxation*, Chicago, University of Chicago Press, 1966; Abba Lerner, *The economics of control*, Nueva York, Macmillan, 1944; Paul Samuelson, *Economics*, Nueva York, McGraw-Hill, 1973 [hay varias ediciones en español; entre ellas: *Curso de economía moderna*, Madrid, Aguilar, 1964; *Economía*, McGraw-Hill /Interamericana de España, S. A., 1986], y "A. P. Lerner at Sixty", en Robert C. Merton (ed.), *Collected scientific papers of Paul A. Samuelson*, Cambridge, Mass., MIT Press, 1972, vol. III.

8 Así, dice Arrow (en "A utilitarian approach", p. 409): "En el análisis utilitarista de la distribución de los ingresos, la igualdad de los ingresos deriva de las condiciones de maximización si se supone, además, que los individuos tienen las mismas funciones de utilidad, cada cual con una utilidad marginal decreciente". Y Samuelson (*Economics*, p. 164 *in fine*) ofrece la siguiente formulación: "Si cada dólar adicional le trae cada vez menos satisfacciones a un hombre, y si los ricos y los pobres son iguales en su capacidad de disfrutar de la satisfacción, se supone que un dólar cobrado a un millonario y entregado a una persona de ingresos medios agrega a la utilidad total más de lo que le resta".

virtud de la falsedad de (a) y de (b), el argumento que vincula la igualdad económica con la maximización de la utilidad agregada es poco sólido.

En lo que a (b) concierne, es evidente que las funciones de utilidad para el dinero de diferentes individuos no se parecen en absoluto. Algunas personas sufren de debilidades o discapacidades físicas, mentales o emocionales, que limitan la satisfacción que pueden obtener. Más aun, incluso dejando a un lado los efectos de las discapacidades específicas, algunas personas simplemente gozan de las cosas más que otras. Todo el mundo sabe que existen, en cualquier nivel de gasto dado, grandes diferencias en las cantidades de utilidad que obtienen las personas que gastan dinero.

En lo que a (a) concierne, existen buenas razones para no esperar una disminución coherente de la utilidad marginal del dinero. En efecto, el hecho de que las utilidades marginales de ciertos bienes tiendan a disminuir no es un principio de razón. Es una generalización psicológica, que se explica mediante consideraciones como, por ejemplo, que las personas tienden muchas veces, después de un tiempo, a sentirse saciadas con lo que han estado consumiendo y que los sentidos suelen perder su frescura después de un estímulo repetido.[9] Todos sabemos que las experiencias de muchas clases se vuelven cada vez más rutinarias e ingratas a medida que se repiten.

Sin embargo, es cuestionable si esto proporciona algún motivo para esperar una disminución de la utilidad marginal del *dinero*, es decir, de cualquier cosa que funcione como un instrumento genérico de intercambio. Aunque la utilidad de todo lo que se puede comprar con dinero disminuyera inevitablemente en el margen, la utilidad del dinero mismo exhibiría, de todas maneras, un patrón diferente. Es muy posible que el dinero estuviera exento del

9 "Luego, según se van añadiendo más unidades al consumo, la utilidad total va aumentando más despacio, pues la capacidad psicológica del sujeto para apreciar el aumento de la cantidad de un bien se hace cada vez menos aguda. El hecho de que los incrementos de la utilidad total van siendo sucesivamente menores los economistas lo describen como sigue: al aumentar la cantidad sucesiva de un bien, disminuye su *utilidad marginal* del bien (o incremento de la utilidad total añadido por la última unidad consumida)" (Samuelson, *Economics*, p. 431 [citamos según la traducción de Aguilar mencionada en nota 7, p. 459.)

fenómeno de la implacable disminución marginal debido a su ilimitada versatilidad proteica. Como lo explican Blum y Kalven:

> Cuando se [...] analiza la cuestión de si el dinero tiene una utilidad decreciente es [...] importante dejar a un lado todas las analogías con respecto a la observación de que los bienes particulares tienen una utilidad decreciente para sus usuarios. No hay necesidad aquí de entrar en el debate sobre el hecho de si en la teoría económica es útil o necesario suponer que los bienes tienen una utilidad decreciente. El dinero es infinitamente versátil. Y aunque todas las cosas que se compran con dinero estén sujetas a una ley de utilidades decrecientes, no se desprende de ello que el dinero mismo lo esté.[10]

A partir de la suposición de que una persona tiende a perder cada vez más el interés en lo que consume a medida que aumenta su consumo de ello, no puede deducirse sencillamente que también debe tender a perder interés en el consumo mismo o en el dinero que hace posible el consumo. Ello se debe a que quizá siempre pueda seguir teniendo a su alcance, por muy cansada que esté de lo que ha estado haciendo, bienes que no probó para comprar y nuevos placeres para disfrutar.

De todas formas, hay muchas cosas de las que las personas no comienzan a cansarse de inmediato desde el principio. De ciertos bienes, en realidad obtienen más utilidad después de un consumo sostenido de la que obtienen al comienzo. Ésta es la situación que se da cada vez que el hecho de apreciar, disfrutar o beneficiarse de otra manera de algo depende de llevar a cabo repetidas pruebas, que sirven como una especie de proceso de *precalentamiento*: por ejemplo, cuando el individuo obtiene relativamente poca gratificación del producto o la experiencia en cuestión hasta que adquiere un gusto especial por él, o se ha vuelto adicto a él, o ha comenzado de algún otro modo a relacionarse con él o a responder a él de manera provechosa. La capacidad de obtener gratificación, entonces, es menor en los primeros puntos de la secuencia de con-

10 W. Blum y H. Kalven, *op. cit.*, pp. 57-58.

sumo que en puntos posteriores. En esos casos, la utilidad marginal no disminuye, sino que aumenta. Quizás sea verdad que una persona termina perdiendo interés en todo, sin excepción. Sin embargo, aunque en cada curva de utilidad haya un punto en que la curva inicia una caída firme e irreversible, no puede suponerse que cada segmento de la curva esté en declive.[11]

III

Cuando la utilidad marginal disminuye, no lo hace a causa de alguna deficiencia de la unidad marginal. Disminuye en virtud de la posición final de esa unidad en una secuencia. Sucede lo mismo cuando la utilidad marginal aumenta: la unidad marginal proporciona una mayor utilidad que las unidades anteriores en virtud del efecto causado por la adquisición o el consumo de esas unidades. Ahora bien, cuando la secuencia consiste en unidades de dinero, lo que corresponde al proceso de precalentamiento –por lo menos, en una característica pertinente e importante– es *ahorrar*. Acumular dinero implica, como el precalentamiento, generar una capacidad para obtener, en algún punto subsiguiente de una secuencia, gratificaciones que no pueden obtenerse antes.

El hecho de que, en ocasiones, puede valer especialmente la pena que una persona ahorre dinero en lugar de gastar cada centavo que obtiene se debe en parte a la incidencia de lo que puede pensarse como *umbrales de utilidad*. Consideremos un artículo

11 Las personas tienden a pensar que, en general, es más importante evitar cierto grado de daño que adquirir un beneficio de magnitud comparable. Puede ser que esto se deba en parte a que suponen que la utilidad disminuye en el margen, puesto que en ese caso, el beneficio adicional tendría menos utilidad que la pérdida correspondiente. Sin embargo, debería notarse que la tendencia a adjudicar un valor inferior a adquirir beneficios que a evitar el daño a veces se revierte: cuando las personas son tan infelices que consideran que "no tienen nada que perder", bien pueden adjudicar un valor más alto a mejorar las cosas que a impedir que se vuelvan peores aun (en una medida comparable). En ese caso, lo que disminuye en el margen no es la utilidad de los beneficios, sino la desutilidad del daño.

con las siguientes características: no es fungible, es la fuente de un tipo de satisfacción nueva e imposible de obtener de otra manera y es tan caro que sólo puede ser adquirido mediante el ahorro. La utilidad de la última unidad de dinero que complete un plan de ahorro para dicho artículo puede ser mayor que la utilidad de cualquiera de las unidades de dinero ahorradas anteriormente. Ése será el caso cuando la utilidad proporcionada por el artículo sea mayor que la suma de las utilidades que podrían obtenerse si el dinero ahorrado fuera gastado a medida que entrara o dividido en partes y usado para comprar otras cosas. En una situación de este tipo, la última unidad de dinero ahorrada permite trasponer un umbral de utilidad.[12]

A veces se argumenta que para cualquier persona racional, en el sentido de que busca maximizar la utilidad generada por sus gastos, la utilidad marginal del dinero necesariamente debe disminuir. Abba Lerner presenta este argumento de la siguiente manera:

> El principio de la utilidad marginal decreciente del ingreso puede derivar de la suposición de que los consumidores gastan sus ingresos en una forma que maximiza la satisfacción que pueden obtener del bien adquirido. Con un ingreso dado, todas

12 En virtud de estos umbrales, una unidad de dinero marginal o incremental puede tener una utilidad conspicuamente mayor que las unidades de dinero que no permiten cruzar un umbral. Así, una persona que usa el dinero que le sobra durante cierto tiempo para introducir alguna mejora intrascendente en su rutina de consumo –quizá comprar carne de una calidad algo superior para la cena de todos los días– puede obtener mucho menos utilidad adicional de esta manera que ahorrando el dinero extra durante algunas semanas para ir a ver una obra de teatro o una ópera maravillosas. El efecto umbral es particularmente esencial para la experiencia de los coleccionistas, que en general sienten mayor satisfacción al obtener el artículo que finalmente completa una colección –cualquiera sea ese artículo– que al obtener cualquiera de los otros artículos de la colección. Obtener el artículo final implica cruzar un umbral de utilidad. Es probable que una colección completa de veinte artículos diferentes, cada uno de los cuales tiene la misma utilidad cuando se lo considera en forma individual, tenga una mayor utilidad para un coleccionista que una colección incompleta que tiene el mismo tamaño, pero que incluye copias. El hecho de que la colección esté completa posee, en sí mismo, utilidad, además de la utilidad proporcionada individualmente por los artículos que conforman la colección.

> las cosas que se compran brindan una satisfacción mayor por el dinero que se gastó en ellas que cualquiera de las otras cosas que podrían haberse comprado en su lugar, pero que no se compraron precisamente por esta razón. De esto se desprende que si el ingreso fuera mayor, las otras cosas que se comprarían con el incremento de ingreso serían cosas que son rechazadas cuando el ingreso es menor porque dan menos satisfacción; y si el ingreso fuera aun mayor, se comprarían cosas aun menos satisfactorias. Cuanto mayor es el ingreso, menos satisfactorias son las cosas adicionales que pueden comprarse con incrementos iguales de ingresos. A esto se refiere el principio de la utilidad marginal decreciente del ingreso.[13]

Lerner trae aquí una comparación entre la utilidad de B(n) –los bienes que el consumidor racional en realidad compra con sus ingresos de n unidades de dinero– y "las otras cosas que podrían haberse comprado en su lugar, pero que no se compraron". Dado que él prefiere comprar B(n) más que las otras cosas, que por hipótesis no cuestan más, el consumidor racional debe considerar que B(n) le ofrece mayor satisfacción que las otras cosas. De aquí Lerner deduce que con n unidades adicionales de dinero el consumidor podría comprar sólo cosas con menor utilidad que B(n); y concluye que, en general, "cuanto mayor es el ingreso, menos satisfactorias son las cosas adicionales que pueden comprarse con incrementos iguales de ingresos". Él sostiene que esta conclusión es equivalente al principio de la utilidad marginal decreciente del ingreso.

Parece evidente que Lerner fracasa en su intento de deducir el principio de esta manera. Una razón de ello es que la cantidad de satisfacción que una persona puede obtener de un bien en particular puede variar de modo considerable según si además posee otros bienes o no. Por tanto, la satisfacción que puede obtenerse a partir de cierto gasto puede ser mayor si ya se hizo algún otro gasto. Supongamos que el costo de una porción de palomitas de maíz es el mismo que el costo de la manteca suficiente para vol-

13 A. Lerner, *op. cit.*, pp. 26-27.

verlas deliciosas; y supongamos que un consumidor racional que adora las palomitas de maíz con manteca siente muy poca satisfacción al comerlas sin ella, pero que, no obstante, las prefiere a la manteca sola. Por tanto, preferirá comprar las palomitas de maíz y no la manteca si debe optar por una de las dos cosas. Supongamos ahora que los ingresos de esta persona aumentan de modo que puede comprar también la manteca. Entonces, podrá conseguir algo de lo que disfruta enormemente: su mayor ingreso hace posible que no sólo compre la manteca además de las palomitas de maíz, sino que disfrute de las palomitas de maíz con manteca. La satisfacción que obtendrá al combinar las palomitas de maíz con la manteca bien puede ser considerablemente mayor que la suma de las satisfacciones que puede obtener de los dos bienes por separado. Aquí, nuevamente, hay un efecto umbral.

En un caso de este tipo, lo que el consumidor racional compra con su mayor ingreso es un bien –B(i)– que, cuando su ingreso era menor, había rechazado en favor de B(n) porque tenerlo solo habría sido menos satisfactorio que tener únicamente B(n). Sin embargo, a pesar de esto, no es verdad que la utilidad del ingreso que usa para comprar B(i) sea menor que la utilidad del ingreso que usaba para comprar B(n). Cuando existe la oportunidad de crear una combinación que es (como las palomitas de maíz con manteca) sinérgica en el sentido de que agregar un bien a otro aumenta la utilidad de cada uno, la utilidad marginal del ingreso no puede declinar aunque la secuencia de artículos marginales –tomando cada uno de estos artículos por separado– exhiba un patrón de utilidades decrecientes.

El argumento de Lerner es defectuoso, además, en virtud de otra consideración. Dado que él habla de "las cosas *adicionales* que pueden comprarse con incrementos iguales de ingresos", evidentemente supone que un consumidor racional usa sus primeras n unidades de dinero para comprar cierto bien y que usa cualquier incremento del ingreso que las supere para comprar algo más. Esto lleva a Lerner a suponer que lo que el consumidor compra cuando su ingreso aumenta en i unidades de dinero (donde i es igual o menor que n) debe ser algo que podría haber comprado y que eligió no comprar cuando su ingreso era sólo de n unidades de dinero.

Esta suposición, sin embargo, no tiene fundamento. Con un ingreso de $(n + i)$ unidades de dinero, el consumidor no necesita usar su dinero para comprar tanto B(n) como B(i). Podría usarlo para comprar algo que costara más que cualquiera de esos bienes, algo que fuera demasiado caro para estar a su disposición antes de que su ingreso aumentara. La cuestión es que si un consumidor racional con un ingreso de n unidades de dinero posterga la compra de un bien en particular hasta que aumente su ingreso, ello no necesariamente significa que "rechazó" comprarlo cuando su ingreso era menor. El bien en cuestión puede haber estado fuera de su alcance en ese momento porque costaba más de n unidades de dinero. Su razón para postergar la compra puede no haber estado en absoluto relacionada con sus expectativas comparativas de satisfacción o con sus preferencias o prioridades.

Se deben considerar dos posibilidades. Supongamos, por un lado, que en vez de comprar B(n) cuando su ingreso es de n unidades de dinero, el consumidor racional ahorra ese dinero hasta que pueda agregarle una suma adicional de i unidades de dinero y luego compra B($n + i$). En este caso, es bastante evidente que su postergación de la compra de B($n + i$) no significa que le adjudique menor valor que a B(n). Por otro lado, supongamos que el consumidor racional se niega a ahorrar para comprar B($n + i$) y que gasta todo el dinero que tiene en B(n). En este caso, también sería un error interpretar que su conducta indica que prefiere B(n) a B($n + i$). Su negativa a ahorrar para B($n + i$) podría deberse simplemente a que considera inútil hacerlo porque no cree que sea razonable confiar en que podrá ahorrar lo suficiente para comprarlo en el momento oportuno.

La utilidad de B($n + i$) puede no sólo ser mayor que la utilidad de B(n) o de B(i) por separado. También puede ser mayor que la suma de sus utilidades. Es decir, cuando adquiere B($n + i$), el consumidor puede estar cruzando un umbral de utilidad. Entonces, la utilidad del incremento i de su ingreso es, en realidad, mayor que la utilidad de las n unidades de dinero a las que se agrega, aunque i equivalga o sea menor que n. En ese caso, el ingreso del consumidor racional no exhibe una utilidad marginal decreciente.

IV

El análisis previo ha establecido que una distribución igualitaria puede no maximizar la utilidad agregada. También puede demostrarse fácilmente que, en virtud de la incidencia de los umbrales de utilidad, hay condiciones bajo las que una distribución igualitaria minimiza la utilidad agregada.[14] Así, supongamos que existe una cantidad suficiente de un recurso en particular (por ejemplo, comida o medicamentos) para permitir que algunos de los miembros de una población –pero no todos– sobrevivan. Digamos que el tamaño de la población es de diez individuos, que una persona necesita por lo menos cinco unidades del recurso en cuestión para vivir y que hay disponibles cuarenta unidades. Si algún miembro de esta población debe sobrevivir, algunos deberán recibir más que otros. Una distribución igual, que le otorga a cada persona cuatro unidades, conduce al peor resultado posible, es decir, todos mueren. Con certeza, en este caso sería moralmente grotesco insistir en la igualdad. Tampoco sería razonable sostener que, en las condiciones especificadas, la mejor situación de algunos se justifica sólo cuando ello es en beneficio de aquellos que están en peor situación. Si los recursos disponibles se usan para salvar a ocho personas, la justificación para hacerlo no es, obviamente, que de alguna manera ello beneficia a los dos miembros de la población a los que se deja morir.

Una distribución igualitaria casi seguramente producirá una pérdida neta de utilidad agregada toda vez que implique que haya menos individuos que tengan, con respecto a alguna necesidad, lo suficiente para mantener la vida; en otras palabras, toda vez que haga falta que un número mayor de individuos esté por debajo del umbral de supervivencia. Por supuesto, también puede ocurrir una pérdida de utilidad aun cuando las circunstancias incluyan un umbral que no separe la vida de la muerte. Distribuir los recursos por igual reducirá la utilidad agregada cada vez que requiera que un número de individuos quede por debajo de *cualquier* umbral de utilidad sin

14 N. Rescher, *Distributive justice*, Indianápolis, Bobbs Merril, 1966, pp. 28-30, analiza condiciones de este tipo.

que se produzca un movimiento compensatorio que ubique a otra cantidad adecuada de individuos por encima de algún umbral.

En condiciones de escasez, entonces, una distribución igualitaria puede ser moralmente inaceptable. Otra respuesta a la escasez es distribuir los recursos disponibles de tal modo que tantas personas como sea posible tengan lo suficiente o, en otras palabras, maximizar el impacto de la suficiencia. Esta alternativa es especialmente apremiante cuando la cantidad de un recurso escaso que constituye lo suficiente coincide con la cantidad que es indispensable para evitar algún daño catastrófico, como en el ejemplo que acabamos de considerar, en el que caer por debajo del umbral de comida suficiente o de medicamentos suficientes significa la muerte. Ahora supongamos que en este ejemplo hay disponibles no sólo cuarenta unidades del recurso vital, sino cuarenta y una. Entonces, maximizar el impacto de la suficiencia proporcionando lo suficiente para cada una de las ocho personas deja una unidad sin distribuir. ¿Qué debería hacerse con esta unidad sobrante?

Se ha mostrado antes que es un error sostener que *donde algunas personas tienen menos que lo suficiente, nadie debería tener más que los demás.* Cuando los recursos son escasos, de modo que es imposible que todos tengan lo suficiente, una distribución igualitaria puede llevar al desastre. Ahora bien, también se podría hacer otra afirmación que parece razonable, pero que también es errónea: *donde algunas personas tienen menos que lo suficiente, nadie debería tener más que lo suficiente.* Si esta afirmación fuera correcta, en el ejemplo que estamos analizando la unidad sobrante debería ir a alguna de las dos personas que no tienen nada. Sin embargo, una unidad adicional del recurso en cuestión no mejorará la condición de una persona que no tiene ninguna. Por hipótesis, esa persona morirá aun con la unidad adicional. Lo que necesita no es una unidad, sino cinco.[15] No puede darse por sentado que la situación de una

15 Podría ser correcto decir que necesita una unidad si existe la posibilidad de que consiga cuatro más, ya que, en ese caso, la unidad sobrante puede considerarse potencialmente un constituyente integral del total de cinco que ubica a la persona del otro lado del umbral de supervivencia. En cambio, si no hay posibilidad de que consiga las cinco, adquirir esa única unidad no contribuye a satisfacer ninguna necesidad.

persona que tiene cierta cantidad de un recurso vital sea necesariamente mejor que la de una persona que tiene una cantidad menor, puesto que la cantidad mayor aun puede ser demasiado pequeña para servir a un propósito útil. Tener la mayor cantidad puede incluso empeorar la situación de una persona. Así, es concebible que en un caso en que una dosis de cinco unidades de algún medicamento sea terapéutica, recibir una dosis de una unidad, en cambio, no sea mejor que no recibir ninguna, sino que en realidad resulte tóxica. Y mientras que una persona que tiene una unidad de comida puede vivir un poco más que alguien sin nada de comida, quizás es peor prolongar el proceso de morirse de hambre durante un breve período que terminar rápidamente con la agonía del hambre.

La afirmación de que nadie debería tener más que lo suficiente mientras alguien tenga menos que lo suficiente resulta, en parte, verosímil, porque se basa en una suposición, verosímil en sí misma, que, sin embargo, es falsa, según la cual dar recursos a las personas que tienen menos que lo suficiente necesariamente significa dar recursos a las personas que los necesitan y, por tanto, mejorar su situación. Sin duda, es más razonable asignar una mayor prioridad a mejorar la condición de aquellos que están necesitados que a mejorar la condición de aquellos que no lo están. Sin embargo, dar recursos adicionales a las personas que tienen menos que lo suficiente de esos recursos, y que, por tanto, están necesitadas, en realidad puede no mejorar su situación. Quienes están debajo de un umbral de utilidad no necesariamente obtienen un beneficio si reciben recursos adicionales que los acercan al umbral. Para ellos, lo crucial es alcanzar el umbral. El mero hecho de acercarse a él puede o bien no ser una ayuda, o bien resultar una desventaja.

De ninguna manera deseo sugerir, por supuesto, que quienes están por debajo de un umbral de utilidad no se beneficien nunca –o lo hagan rara vez– del hecho de acercarse a él. Ciertamente pueden beneficiarse de ello, ya sea porque aumenta la probabilidad de que alcancen el umbral, ya porque, aparte de la importancia del umbral, los recursos adicionales proporcionan importantes incrementos de utilidad. Después de todo, un coleccionista puede disfrutar de ampliar su colección aunque sepa que no tiene ninguna posibilidad de completarla. Lo que deseo sostener es sólo que los recursos adicio-

nales no necesariamente benefician a quienes tienen menos que lo suficiente. Los agregados pueden ser demasiado pequeños como para tener algún impacto. Por tanto, puede ser moralmente aceptable que algunos tengan más que lo suficiente de cierto recurso mientras que otros tienen menos que lo suficiente de ese mismo recurso.

V

Con mucha frecuencia, la defensa del igualitarismo se basa menos en un argumento que en una supuesta intuición moral: la desigualdad económica, considerada como tal, parece ser simplemente algo incorrecto. A muchas personas les parece indudablemente obvio que, tomado sólo por sí mismo, el hecho de que algunos disfruten de beneficios económicos mayores que aquellos de los que disfrutan otros es moralmente ofensivo. Sospecho, sin embargo, que en muchos casos quienes manifiestan tener esta intuición con respecto a las manifestaciones de la desigualdad no están respondiendo a la desigualdad, sino a otro rasgo de las situaciones a las que se enfrentan. Creo que lo que por intuición les parece moralmente objetable, en los tipos de situaciones citados habitualmente como ejemplos de desigualdad económica, no es el hecho de que algunos de los individuos en esas situaciones tengan *menos* dinero que otros, sino el hecho de que aquellos que tienen menos tienen *demasiado poco*.

Cuando pensamos en las personas cuya situación es bastante peor que la nuestra, es muy común que nos encontremos moralmente perturbados por sus circunstancias. Lo que nos toca directamente en casos de este tipo, sin embargo, no es una discrepancia cuantitativa, sino una condición cualitativa; no el hecho de que los recursos económicos de aquellos cuya situación es peor sean *de inferior magnitud* que los nuestros, sino el hecho de que estas personas sean tan *pobres*. Las meras diferencias en la cantidad de dinero que tienen las personas no son, en sí mismas, angustiantes. Después de todo, solemos no conmovernos por las desigualdades entre los acomodados y los ricos; nuestra concien-

cia de que la situación de los primeros es bastante peor que la de los segundos no nos perturba en absoluto desde el punto de vista moral. Y si creemos que la vida de alguna persona es rica y satisfactoria, que ella misma está realmente conforme con su situación económica y que no sufre resentimientos ni pesares que podrían aliviarse con más dinero, en general no estamos muy interesados –desde un punto de vista moral– en comparar la cantidad de dinero que tiene con la cantidad que poseen otros. Las discrepancias económicas en casos de este tipo no nos parecen asuntos de un interés moral importante. El hecho de que algunas personas tengan mucho menos que otras no es moralmente perturbador cuando está claro que tienen bastante.

Parece claro que el igualitarismo y la doctrina de la suficiencia son lógicamente independientes: no se puede suponer que las consideraciones que apoyan el primero proporcionan apoyo también para la segunda. Sin embargo, los defensores del igualitarismo creen, muchas veces, que han ofrecido fundamentos para su posición cuando, en realidad, lo que han ofrecido sólo es un apoyo pertinente para la doctrina de la suficiencia. Así, muchas veces, cuando intentan ganar aceptación para el igualitarismo, llaman la atención hacia las disparidades entre las condiciones de vida características de los ricos y aquéllas características de los pobres. Ahora bien, es innegable que contemplar estas disparidades provoca con frecuencia la convicción de que sería moralmente deseable redistribuir los recursos disponibles a fin de mejorar las circunstancias de los pobres. Y, por supuesto, esto daría lugar a un mayor grado de igualdad económica. Sin embargo, el carácter irrefutable de la apelación moral a mejorar la situación de los pobres asignándoles recursos tomados de aquellos cuya situación es mejor ni siquiera tiende a mostrar que el igualitarismo es, como ideal moral, irrefutable de un modo similar. Demostrar que la pobreza es fuertemente indeseable no contribuye en absoluto a demostrar lo mismo de la igualdad. Lo que hace que alguien sea pobre en el sentido moral pertinente –en el que la pobreza se entiende como una situación ante la que naturalmente retrocedemos– no es que sus bienes económicos sean sólo de menor magnitud que los de los demás.

Ronald Dworkin proporciona un típico ejemplo de esta confusión. Dworkin afirma que el ideal de igualdad económica requiere que "ningún ciudadano tenga menos que una parte proporcionalmente igual de los recursos de una comunidad a fin de que los demás tengan más de lo que él carece".[16] Sin embargo, para respaldar su afirmación de que los Estados Unidos no alcanzan ahora este ideal, se refiere a circunstancias que no son una prueba específica de la desigualdad, sino de la pobreza:

> Creo que es evidente que los Estados Unidos actualmente no alcanzan [el ideal de la igualdad]. Una minoría considerable de estadounidenses está crónicamente desempleada o recibe salarios por debajo de cualquier *línea de pobreza* realista, está discapacitada de diversas maneras o agobiada con necesidades especiales; y la mayoría de estas personas harían el trabajo necesario para ganarse un salario decente si tuvieran la oportunidad y la capacidad (p. 208).

Lo que le preocupa principalmente a Dworkin –lo que en realidad considera de suma importancia moral– no es que nuestra sociedad permita una situación en la que una minoría considerable de estadounidenses tenga *porciones más pequeñas* que otros de los recursos que, al parecer, él supone que deberían estar disponibles para todos. Más bien, su preocupación es que los miembros de esta minoría *no ganan salarios decentes.*

La fuerza de la denuncia de Dworkin no deriva del argumento de que nuestra sociedad no provee a algunos individuos lo mismo que a otros, sino de un argumento muy diferente, a saber, que nuestra sociedad no provee a cada individuo "la oportunidad de desarrollarse y de llevar una vida que pueda considerar valiosa tanto para sí mismo como para [la comunidad]" (p. 211). Dworkin se siente consternado, fundamentalmente, no por la evidencia de que los Estados Unidos permiten la desigualdad económica,

16 R. Dworkin, "Why liberals should care about equality", en su *A matter of principle*, Cambridge, Mass., Harvard University Press, 1985, p. 206. Los números entre paréntesis en los párrafos que siguen se refieren a páginas de este ensayo.

sino por la evidencia de que no garantizan que todos tengan lo suficiente para llevar "una vida de elección y de valor" (p. 212); en otras palabras, que no cumplen para todos el ideal de la suficiencia. Lo que le molesta no es que estén generalizadas ciertas relaciones cuantitativas, sino que prevalezcan ciertas situaciones cualitativas. Se preocupa principalmente por el valor de la vida de las personas, pero se representa a sí mismo, en forma errónea, como alguien que se preocupa más que nada por la magnitud relativa de sus bienes económicos.

Me parece que las conocidas discrepancias entre los principios que profesan los igualitaristas y la forma en que, en general, conducen su propia vida confirman mi sugerencia de que las situaciones que involucran la desigualdad son moralmente inquietantes sólo en la medida en que violan el ideal de la suficiencia. El problema no es que algunos igualitaristas acepten hipócritamente los altos ingresos y las oportunidades especiales para las que, según las teorías morales que profesan, no existe justificación. Es, en cambio, que muchos igualitaristas (incluyendo a muchos defensores académicos de la doctrina) no están realmente preocupados por la cuestión de si su situación económica es la misma que la de los demás. Creen que ellos mismos tienen dinero apenas suficiente para lo que es importante para ellos y, por tanto, no están demasiado preocupados por el hecho de que algunas personas son considerablemente más ricas que ellos. Sin duda, muchos igualitaristas considerarían que es algo bastante ruin y censurable, con respecto a su propia vida, preocuparse por las comparaciones económicas de ese tipo. Además, a pesar de las consecuencias de las doctrinas a las que instan a adherirse, se sentirían consternados si sus hijos crecieran con esas preocupaciones.

VI

El error fundamental del igualitarismo radica en suponer que es moralmente importante el hecho de si una persona tiene menos que otra sin que importe cuánto tiene cada una de ellas. Este error

se debe, en parte, a la falsa suposición de que alguien cuya situación económica es peor tiene necesidades insatisfechas más importantes que alguien cuya situación es mejor. De hecho, las necesidades moralmente importantes de ambos individuos pueden estar o completamente satisfechas o igualmente insatisfechas. La cuestión de si una persona tiene más dinero que otra es totalmente extrínseca. Tiene que ver con la relación entre los bienes económicos respectivos de las dos personas, lo cual no sólo es independiente de la cantidad de sus bienes y de la cantidad de satisfacción que pueden obtener de ellos, sino también de las actitudes de estas personas hacia esos niveles de bienes y de satisfacción. La comparación económica no responde a la pregunta de si alguna de las personas comparadas tiene alguna necesidad moral importante insatisfecha, ni tampoco a la de si alguna está conforme con lo que tiene.

Esta falla del igualitarismo aparece claramente en el desarrollo que Thomas Nagel hace de la doctrina. De acuerdo con Nagel:

> La característica esencial de un sistema igualitario de prioridades es que considera que las mejoras del bienestar de aquellos que están en peor situación son más urgentes que las mejoras del bienestar de aquellos que están mejor. [...] Lo que hace que un sistema sea igualitario es la prioridad que da al reclamo de quienes están [...] en los niveles inferiores. [...] Cada individuo con una demanda más urgente tiene prioridad [...] sobre cada individuo con una demanda menos urgente.[17]

Al comentar el Principio de las Diferencias de Rawls, al que adhiere, Nagel dice que este principio "establece un orden de prioridades entre las necesidades y da preferencia a las más urgentes".[18] Sin embargo, la preferencia que en realidad asigna el Principio de las Diferencias no está a favor de aquellos cuyas necesidades son más urgentes; está a favor de aquellos cuya situación es peor. Nagel está suponiendo, sin proporcionar ningún fundamento para ello, que los individuos en peor situación tienen necesidades urgentes.

17 T. Nagel, *op. cit.*, p. 118.
18 *Ibid.*, p. 117.

En la mayoría de las sociedades, las personas que están en los niveles económicos inferiores son, sin duda, en extremo pobres y tienen, de hecho, necesidades urgentes. Sin embargo, esta relación entre una posición económica baja y necesidades urgentes es totalmente contingente. Puede establecerse sólo sobre la base de datos empíricos. No existe una conexión conceptual necesaria entre la posición económica relativa de una persona y la cuestión de si tiene necesidades de algún grado de urgencia.[19]

Es posible que aquellos que están en peor situación no tengan necesidades o demandas más urgentes que aquellos que están mejor, porque es posible que no tengan necesidades o demandas urgentes en absoluto. La noción de *urgencia* tiene que ver con lo que es *importante*. No es correcto interpretar que las necesidades o los intereses triviales –que no tienen ninguna relación significativa con la calidad de vida de una persona o con su disposición para estar conforme con ella– son urgentes en ninguna medida o respaldan la clase de demandas moralmente exigentes a que da lugar la urgencia genuina. Más aun, a partir del hecho de que una persona está en los niveles inferiores de un orden económico, ni siquiera puede inferirse que tenga *alguna* necesidad o demanda insatisfechas. Después de todo, es posible que las condiciones de los niveles inferiores sean bastante buenas; el hecho de que sean las peores no implica en sí mismo que sean malas o, de alguna manera, incompatibles con vidas ricamente satisfactorias y placenteras.

Nagel sostiene que lo que subyace a la apelación a la igualdad es un "ideal de aceptabilidad para cada individuo".[20] En su trabajo, este ideal implica que una persona razonable debería aceptar las desviaciones respecto de la igualdad sólo si resultan en su benefi-

19 Me opongo a la afirmación de que cuando se trata de justificar los intentos por mejorar las circunstancias de aquellos cuya situación económica es la peor, una buena razón para hacer el intento es que es moralmente importante que las personas alcancen la mayor igualdad posible con respecto al dinero. La única razón moral apremiante para tratar de que los que peor están estén mejor es, a mi juicio, que sus vidas son, en cierta medida, vidas malas. El hecho de que algunas personas tengan más que el dinero suficiente sugiere una forma en la que podría disponerse que quienes tienen menos que lo suficiente obtengan más, pero no es una buena razón en sí misma para la redistribución.

20 T. Nagel, *op. cit.*, p. 123.

cio, en el sentido de que su situación sería peor sin ellas. Sin embargo, una persona razonable bien podría considerar aceptable una distribución desigual aunque no supusiera que cualquier otra distribución la beneficiaría menos. Ello se debe a que podría creer que la distribución desigual le proporcionó lo suficiente; y sería razonable que estuviera inequívocamente conforme con eso, sin preocuparse por la posibilidad de que algún otro arreglo le proporcionara más. Es gratuito suponer que todas las personas razonables deben estar buscando maximizar los beneficios que pueden obtener, lo cual requeriría, en un sentido, que estuvieran continuamente interesadas en mejorar sus vidas o abiertas a esa posibilidad. Cierta desviación respecto de la igualdad podría no resultar *en* beneficio de alguien, porque podría suceder que, de hecho, su situación fuera mejor sin ella. Sin embargo, siempre que no *provoque un conflicto* con su interés, al obstruir su oportunidad de llevar la clase de vida que, para él, es importante llevar, la desviación respecto de la igualdad puede ser muy aceptable. A fin de estar completamente satisfecha con cierta situación, no es necesario que una persona razonable suponga que no existe otra situación en la que estaría mejor.[21]

Nagel ejemplifica su tesis acerca de la apelación moral a la igualdad considerando una familia con dos hijos, uno de los cuales es "normal y bastante feliz" mientras que el otro "sufre de una dolorosa discapacidad".[22] Si esta familia se instalara en la ciudad, el hijo discapacitado se beneficiaría de las oportunidades médicas y educativas que no existen en las afueras, pero el hijo sano se divertiría menos. Si la familia se instalara en las afueras, en cambio, el hijo discapacitado estaría en desventaja, pero el hijo sano se divertiría más. Nagel estipula que sería mayor la ventaja para el hijo sano si se instalaran en las afueras que la que el hijo discapacitado obtendría si se instalaran en la ciudad: en la ciudad, el hijo sano encontraría la vida indudablemente desagradable, mientras que el hijo discapacitado no sería feliz, "sino sólo menos infeliz".

21 Para un análisis más detallado, véase la sección VII más adelante.

22 Las citas del análisis de Nagel de esta ilustración son de su *Mortal questions*, pp. 123-124.

Dadas estas consideraciones, la decisión igualitaria es instalarse en la ciudad, puesto que "es más urgente beneficiar al hijo [discapacitado] aunque el beneficio que podamos brindarle sea menor que el beneficio que podemos darle al hijo [sano]". Nagel explica que esta evaluación acerca de la mayor urgencia de beneficiar al hijo discapacitado "depende de la peor posición del hijo [discapacitado]. Una mejora de su situación es más importante que una mejora igual o algo mayor de la situación del hijo [normal]". Sin embargo, me parece que el análisis que Nagel hace de este asunto es defectuoso debido a un error similar al que le atribuí antes a Dworkin. El hecho de que sea preferible ayudar al hijo discapacitado no se debe, como afirma Nagel, a que la situación de este niño sea peor que la del otro. Se debe al hecho de que este niño, y no el otro, sufre de una discapacidad dolorosa. La demanda del hijo discapacitado es importante porque su situación es *mala* –significativamente indeseable– y no sólo porque es *peor* que la de su hermano.

Por supuesto, esto no implica que la evaluación de Nagel acerca de lo que debería hacer la familia sea mala. Con toda certeza, rechazar el igualitarismo no significa sostener que siempre sea obligatorio sencillamente maximizar los beneficios y que, por tanto, la familia debería instalarse en las afueras de la ciudad porque el hijo normal saldría ganando más con esa decisión de lo que ganaría el hijo discapacitado si se instalaran en la ciudad. Sin embargo, la base más convincente para el juicio de Nagel en favor del hijo discapacitado no tiene nada que ver con la supuesta urgencia de proporcionar a unos tanto como a otros. Pertenece más bien a la urgencia de las necesidades de las personas que no tienen lo suficiente.[23]

VII

¿Qué significa, en este contexto, que una persona tenga lo suficiente? Una cosa que podría significar es que más sería demasiado:

23 Por supuesto, la cuestión de la igualdad o la suficiencia que plantea el ejemplo de Nagel no concierne a la distribución de *dinero*.

una mayor cantidad haría que la vida de la persona fuera desagradable, o sería perjudicial o, de alguna otra forma, sería poco grata. Muchas veces, esto es lo que piensan las personas cuando dicen cosas como "¡Ya he tenido suficiente!" o "¡Suficiente con eso!". La idea que transmiten las expresiones como éstas es que *se ha llegado a un límite*, más allá del cual no es conveniente continuar. Por otro lado, la afirmación de que una persona tiene lo suficiente también puede significar sólo que *se ha alcanzado cierto requerimiento o estándar*, sin que ello implique que una mayor cantidad sería algo malo. Con frecuencia, esto es a lo que una persona se refiere cuando dice algo así como "Esto debería ser suficiente". Los enunciados como éste determinan que la cantidad indicada es suficiente al tiempo que dejan abierta la posibilidad de que una cantidad superior también podría ser aceptable.

En la doctrina de la suficiencia, el uso de la noción de *suficiente* tiene que ver con *alcanzar un estándar* más que con *llegar a un límite*. Decir que una persona tiene suficiente dinero significa que está conforme o que es razonable que esté conforme con el hecho de no tener más dinero del que tiene. Y, a su vez, decir esto es decir algo como lo siguiente: la persona no considera (o no puede hacerlo razonablemente) que cualquier cosa (si es que hay algo) que sea insatisfactoria o angustiante en su vida se debe a que tiene demasiado poco dinero. En otras palabras, si una persona está conforme (o razonablemente debería estarlo) con la cantidad de dinero que tiene, entonces, en la medida en que sea infeliz con su vida o tenga razones para serlo, no supone (o razonablemente no puede hacerlo) que más dinero le permitiría –ya como condición suficiente, ya necesaria– ser (o tener razón para ser) significativamente menos infeliz con ella.[24]

Es esencial entender que tener suficiente dinero es diferente de tener sólo lo suficiente para vivir o lo suficiente para que nuestra vida sea marginalmente tolerable. Las personas en general no están conformes con el hecho de vivir en el límite. La cuestión central de

24 Dentro de los límites de mi análisis, no importa qué punto de vista se tome con respecto a la relevante cuestión de si lo que cuenta es *la actitud que una persona en realidad tiene* o *la actitud que sería razonable que tuviera*. A fin de ser breve, en adelante omitiré referirme a la última alternativa.

la doctrina de la suficiencia no es que la única consideración moralmente importante con respecto a la distribución del dinero es la cuestión de si las personas tienen lo suficiente para evitar la miseria económica. Una persona de la que podría decirse, en forma natural y apropiada, que tiene apenas lo suficiente, en realidad no tiene para nada lo suficiente, según el estándar invocado en la doctrina de la suficiencia.

Existen dos tipos distintos de circunstancias en las que la cantidad de dinero que tiene una persona es suficiente; es decir, en las que más dinero no la hará significativamente menos infeliz. Por un lado, puede ser que la persona no esté angustiada ni insatisfecha con su vida. Por el otro, puede ser que, a pesar de que la persona sea infeliz con su vida, las dificultades que causan su infelicidad no pudieran ser aliviadas con más dinero. Las circunstancias de este segundo tipo imperan cuando el problema con la vida de la persona tiene que ver con bienes que no son de índole económica, tales como el amor, la sensación de que la vida tiene sentido, la satisfacción con el propio carácter, etc. Éstos son bienes que el dinero no puede comprar; más aun, son bienes para los cuales ninguna de las cosas que el dinero puede comprar puede servir siquiera como un sustituto adecuado. A veces, ciertamente los bienes no económicos pueden ser obtenidos o disfrutados sólo (o con mayor facilidad) por alguien que tiene cierta cantidad de dinero. Sin embargo, la persona que está angustiada con su vida y, a la vez, conforme con su situación económica puede tener ya esa cantidad de dinero.

Es posible que alguien que está conforme con la cantidad de dinero que tiene también esté conforme con una cantidad de dinero aun mayor. Dado que tener dinero suficiente no significa estar en el límite más allá del cual más dinero resultaría necesariamente indeseable, sería un error suponer que para una persona que ya tiene dinero suficiente la utilidad marginal del dinero debe ser negativa o cero. Si bien, por hipótesis, esta persona no está angustiada por su vida, ya que no le hace falta nada que podría obtener con más dinero, sigue siendo posible que disfrutara del hecho de tener algunas de esas cosas. No la harían menos infeliz, como tampoco alterarían su actitud hacia su vida ni el grado de conformidad con ella en forma alguna, pero podrían darle placer. Si así fuera, en-

tonces su vida, en este aspecto, sería mejor con más dinero que sin él. Por tanto, la utilidad marginal del dinero seguiría siendo para ella positiva.

Decir que una persona está conforme con la cantidad de dinero que tiene no implica, entonces, que no tendría ningún sentido que tuviera más. Así, alguien con dinero suficiente podría estar muy *dispuesto* a aceptar mayores beneficios económicos. De hecho, podría estar *contento* de recibirlos. En efecto, a partir de la suposición de que una persona está conforme con la cantidad de dinero que tiene, no puede deducirse siquiera que no *preferiría* tener más. Y hasta es posible que en realidad estuviera dispuesta a *sacrificar* ciertas cosas que valora (por ejemplo, parte de su tiempo libre) para obtener más dinero.

Ahora bien, ¿cómo puede ser compatible todo esto con la declaración de que la persona está conforme con lo que tiene? ¿Qué *imposibilita* estar conforme con una cantidad dada de dinero, si no es el hecho de preferir más dinero, estar dispuesto a tener más, estar contento de recibirlo o, incluso, aceptar hacer sacrificios para tener más? Imposibilita que la persona tenga un *interés activo* en obtener más. Una persona conforme considera que tener más dinero *no es esencial* para estar satisfecha con su vida. El hecho de que esté conforme es coherente con el de que reconozca que sus circunstancias económicas podrían mejorar y que su vida podría, en consecuencia, ser mejor de lo que es. Sin embargo, esta posibilidad no es importante para ella. Simplemente no está interesada en mejorar su situación, en lo que al dinero se refiere. Su atención y su interés no están vívidamente comprometidos con los beneficios que estarían disponibles para ella si tuviera más dinero. Simplemente, no reacciona mucho a su atractivo. No suscitan en ella ningún interés particularmente ansioso o inquieto, aunque reconoce que disfrutaría de otros beneficios si los recibiera.

En cualquier caso, supongamos que el nivel de satisfacción que le permiten obtener sus circunstancias económicas actuales es lo bastante alto para satisfacer sus expectativas de vida. Esto no depende, en esencia, de cuánta utilidad o satisfacción le proveen sus diversas actividades y experiencias. Más bien, decididamente depende de su actitud hacia la circunstancia de recibir todo eso. Las

experiencias satisfactorias que tiene una persona son una cosa. Otra muy diferente es la cuestión de si está satisfecha con que su vida incluya sólo esas satisfacciones. Si bien es posible que otras circunstancias viables le proporcionaran mayor satisfacción, puede ser que esté completamente satisfecha con el grado de satisfacción del que ahora disfruta. Aunque sabe que podría obtener una mayor satisfacción en general, no experimenta la incomodidad o la ambición que la inclinarían a buscarla. Algunas personas sienten que sus vidas son lo suficientemente buenas, y no les resulta importante la cuestión de si son lo mejor que pueden ser.

El hecho de que una persona carezca de un interés activo en obtener algo no significa, por supuesto, que prefiera no tenerlo. Éste es el motivo por el que la persona conforme puede, sin que haya nada incoherente en ello, aceptar o recibir de buen grado mejoras en su situación, y la razón por la que incluso puede estar dispuesta a incurrir en pequeños gastos para mejorarla. El hecho de que esté conforme sólo significa que la posibilidad de mejorar su situación no es *importante* para ella. Sólo implica, en otras palabras, que no está enojada con sus circunstancias, que no está ansiosa por mejorarlas ni decidida a hacerlo, y que no se desvive por tomar ninguna iniciativa importante para ello.

Puede parecer que no existe una base razonable para aceptar menos satisfacción cuando se podría tener más; que, por tanto, la racionalidad misma implica la maximización y, como consecuencia, que una persona que se niega a maximizar la cantidad de satisfacción que hay en su vida no actúa racionalmente. Esa persona no puede, por supuesto, ofrecer como razón para rehusarse a buscar una mayor satisfacción el que los costos de su búsqueda son demasiado altos, puesto que si ésa fuera su razón, entonces, después de todo, claramente estaría tratando de maximizar la satisfacción. Sin embargo, ¿qué otra buena razón podría tener para dejar pasar una oportunidad de obtener más satisfacción? De hecho, es posible que tenga una muy buena razón para esto: a saber, *que está satisfecha con la cantidad de satisfacción que ya tiene.* Estar satisfecho con el estado de las cosas es, sin duda, una excelente razón para no tener mayor interés en cambiarlo. Por tanto, una persona que, en efecto, está satisfecha con su vida tal como es no

puede ser criticada aduciendo que no tiene una buena razón para rehusarse a mejorarla.

Incluso podría estar abierta a la crítica sobre la base de que *no debería* estar satisfecha, es decir que, de alguna manera, es poco razonable o impropio, o bien incorrecto que se sienta satisfecha con menos satisfacción de la que podría obtener. Sin embargo, ¿sobre qué base podría justificarse *esta* crítica? ¿Acaso existe alguna razón decisiva para insistir en que una persona debería ser tan difícil de satisfacer? Supongamos que un hombre ama, profunda y felizmente, a una mujer que es por completo respetable. En general, no criticamos al hombre en ese caso sólo porque pensemos que podría haber logrado algo mejor. Más aun, el hecho de que sintamos que sería inapropiado criticarlo por esa razón no necesariamente se debe sólo a una creencia de que tratar de buscar a una mujer más deseable o más meritoria podría terminar costándole más de lo que valdría la pena. Más bien, puede reflejar nuestro reconocimiento de que el deseo de ser felices, estar conformes o satisfechos con la vida es un deseo de alcanzar un grado de satisfacción satisfactorio y no equivale, en sí mismo, a un deseo de que la satisfacción sea mayor.

El hecho de estar satisfecho con una situación determinada no equivale a preferirla a todas las demás. Si una persona se enfrenta a una elección entre menos o más de algo deseable, entonces, sin duda sería irracional que prefiriera menos a que prefiriera más. Sin embargo, una persona puede estar satisfecha sin haber hecho en absoluto estas comparaciones. Tampoco es necesariamente irracional o poco razonable que una persona omita o rechace hacer comparaciones entre su propia situación y las posibles alternativas. Ello no sólo se debe a que hacer comparaciones puede resultar demasiado costoso. También se debe a que si alguien está satisfecho con las cosas como son, tal vez no tenga motivo para considerar de qué otra manera podrían ser.[25]

La conformidad podría ser una función de la monotonía o el retraimiento excesivos. El hecho de que una persona esté libre tanto de resentimiento como de ambición puede deberse a que posee un

25 Piénsese en el sensato adagio: "Si no está roto, no lo arregles".

carácter servil o a que su vitalidad se ve acallada por una suerte de lasitud negligente. Es posible que alguien esté conforme, por así decirlo, simplemente por defecto. No obstante, tal vez una persona que está conforme con recursos que proporcionan menos utilidad de la que podría obtener no sea ni irresponsable ni indolente, ni posea una imaginación deficiente. Por el contrario, su decisión de estar conforme con esos recursos –en otras palabras, de adoptar una actitud de aceptación dispuesta hacia el hecho de que tiene lo que tiene– puede estar basada en una evaluación concienzuda, inteligente y profunda de las circunstancias de su vida.

No es imprescindible que dicha evaluación incluya una comparación *extrínseca* de las circunstancias de la persona con alternativas a las que razonablemente podría aspirar, como debería hacerlo si la conformidad sólo fuera razonable cuando se basara en la evaluación de que se ha maximizado el disfrute de posibles beneficios. Si alguien está menos interesado en el hecho de si sus circunstancias le permiten vivir lo mejor posible que en el hecho de si le permiten vivir en forma satisfactoria, bien podría dedicar su evaluación por completo a una apreciación *intrínseca* de su vida. Entonces, tal vez reconozca que sus circunstancias lo conducen a no sentir resentimiento ni pesar, ni verse atraído por el cambio y que, sobre la base de su comprensión de sí mismo y de lo que es importante para él, accede con aprobación a su verdadera disposición a estar conforme con el estado de las cosas. En ese caso, no se trata de que rechace la posibilidad de mejorar sus circunstancias porque cree que, en verdad, que no hay nada que ganar tratando de mejorarlas. Más bien se trata de que esta posibilidad, por viable que pueda ser, de hecho no provoca su atención activa ni suscita en él ningún vívido interés.[26]

26 Las personas suelen adaptar sus deseos a sus circunstancias. Existe el peligro de que el mero desaliento o el interés por evitar la frustración y el conflicto las conduzca a conformarse con demasiado poco. Por cierto, no puede suponerse que la vida de alguien es genuinamente satisfactoria, o que es razonable que la persona esté satisfecha con ella sólo porque no se queja. Por otro lado, tampoco puede suponerse que cuando una persona ha acomodado sus deseos a sus circunstancias, eso sea, en sí mismo, la prueba de que algo salió mal.

APÉNDICE

El igualitarismo económico es una doctrina áridamente formalista. La cantidad de dinero que sus seguidores quieren para sí mismos y para otros se calcula sin considerar las características o las circunstancias personales de nadie. En esta formalidad, los igualitaristas parecen personas que desean ser lo más ricas posible, pero que no tienen idea de qué hacer con su riqueza. En ningún caso las ambiciones individuales, en lo que al dinero se refiere, se limitan o se miden de acuerdo con la comprensión de las metas para las que la persona desea que sirva su dinero o de la importancia que estas metas tienen para ella.

El deseo de riqueza ilimitada es fetichista en la medida en que refleja, con respecto a un *medio*, una actitud –es decir, desear algo por sí mismo– que es apropiada sólo con respecto a un *fin*. Me parece que la actitud que adopta John Rawls hacia lo que llama *bienes primarios* ("derechos y libertades, oportunidades y poderes, ingreso y riqueza")[27] tiende al fetichismo en este sentido. Los bienes primarios son "medios para todo propósito" –explica Rawls–, que las personas necesitan más allá de otras cosas que quieran: "Los planes difieren, dado que las capacidades, las circunstancias y los deseos individuales difieren [...]; sin embargo, cualquiera sea nuestro sistema de fines, los bienes primarios son un medio necesario" (p. 93). A pesar del hecho de que no identifica los bienes primarios como fines sino como medios, Rawls considera racional que una persona quiera tantos de ellos como sea posible. Así, dice:

> Independientemente de cuáles sean en detalle los planes racionales de un individuo, se supone que hay diversas cosas de las que preferiría más en vez de menos. [...] Mientras que las personas que están en la posición original no saben cuál es su concepción del bien, sí saben, supongo, que prefieren más bienes primarios en vez de menos (pp. 92-93).

27 J. Rawls, *A theory of justice,* Cambridge, Mass., Harvard University Press, 1971, p. 92 [trad. esp.: *Teoría de la justicia*, México, Fondo de Cultura Económica, 1979].

La suposición de que siempre debe ser mejor tener más bienes primarios que menos implica que la utilidad marginal de una cantidad adicional de un bien primario es invariablemente mayor que su costo. En otras palabras, implica que la mayor ventaja para un individuo de poseer una mayor cantidad de bienes primarios nunca es superada por mayores desventajas, discapacidades o cargas que les correspondan.

No obstante, esto parece muy poco razonable. Aparte de cualquier otra consideración, poseer más de un bien primario puede requerir que un individuo responsable invierta más tiempo y esfuerzo para ocuparse de él y para tomar decisiones respecto de su uso. Para muchas personas, estas actividades son intrínsecamente poco atractivas; y en general, también suponen cierta carga de ansiedad y un grado de distracción respecto de otros intereses. Ciertamente, no debe darse por sentado que los mayores costos de esta clase nunca puedan superar cualquier beneficio aumentado que proporcionaría una mayor cantidad correspondiente de algún bien primario.

Los individuos ubicados en la posición original están detrás de un velo de ignorancia. No conocen sus propias concepciones del bien o sus propios planes de vida. Por tanto, puede parecerles racional elegir poseer bienes primarios en cantidades ilimitadas: dado que no saben para qué prepararse, quizá sería mejor que estuvieran preparados para cualquier cosa. Sin embargo, incluso en la posición original es posible que las personas aprecien que, en cierto punto, el costo de bienes primarios adicionales podría exceder los beneficios que esos bienes proporcionan. Es verdad que un individuo que está detrás de un velo de ignorancia no puede saber exactamente en qué punto encontraría que un aumento de su provisión de bienes primarios cuesta más de lo que vale. Sin embargo, su ignorancia de la exacta ubicación de ese punto no justifica que actúe como si ese punto no existiera en absoluto. No obstante, actúa precisamente de esa manera si elige que la cantidad de bienes primarios que posee sea ilimitada.

Rawls reconoce que las cantidades adicionales de bienes primarios pueden ser, para ciertos individuos, más caras de lo que valen. Sin embargo, en su opinión, esto no invalida la suposición de que

es racional que todos los que están en la posición original quieran todos los bienes primarios que puedan obtener. Él lo explica de la siguiente manera:

> Postulo que ellas [es decir, las personas que están en la posición original] suponen que preferirían más bienes sociales primarios que menos de ellos. Por supuesto, una vez que se retira el velo de la ignorancia, puede resultar que algunas de ellas, por razones religiosas o de otra índole, en efecto, no quieran más de estos bienes. Sin embargo, desde el punto de vista de la posición original, es racional que las partes supongan que sí quieren que su porción sea mayor, ya que, de todas formas, no están obligadas a aceptar más si no desean hacerlo, como tampoco una persona sufre con una mayor libertad (pp. 142-143).

No me parece que este argumento sea convincente. Descuida el hecho de que prescindir de bienes primarios disponibles o rechazarlos es, en sí misma, una acción que puede implicar costos significativos. Pueden hacer falta deliberaciones y cálculos onerosos para que una persona determine si vale la pena tener un incremento de algún bien primario; y tomar decisiones de este tipo puede suponer responsabilidades y riesgos en virtud de los cuales la persona experimenta una considerable angustia. Además, ¿cuál es la base para afirmar que nadie sufre con una mayor libertad? Al parecer, en una variedad de circunstancias, puede ser razonable que las personas prefieran tener menos alternativas entre las cuales elegir en vez de más. Ciertamente, la libertad, como todo lo demás, tiene sus costos. Es un error suponer que la vida de una persona invariablemente mejora, o que no puede empeorar, cuando aumentan sus opciones.[28]

28 Para un análisis pertinente de este tema, véase Gerald Dworkin, "Is more choice better that less?", en P. French, T. Uchling y H. Wettstein (comps.), *Midwest studies in philosophy* VII, Minneapolis, 1982.

12
Identificación e incondicionalidad

I

La frase "el problema de la mente y el cuerpo" es tan clara, y su papel en el discurso filosófico está tan bien establecido, que oponerse a su uso sería simplemente necio. No obstante, el modo en que se usa *es* algo anacrónico. El conocido problema al que hace referencia la frase concierne a la relación entre el cuerpo de una criatura y el hecho de que la criatura sea consciente. Un nombre más apropiado sería, por tanto, "el problema de la conciencia y el cuerpo", pues ya no es verosímil equiparar el ámbito de los fenómenos conscientes –como hizo Descartes– con el ámbito de la mente. Esto no se debe sólo a que el psicoanálisis haya tornado imprescindible la noción de sentimientos y pensamientos inconscientes. Otras teorías psicológicas capitales también han encontrado útil interpretar que la distinción entre lo mental y lo que no es mental es mucho más amplia que la que existe entre situaciones en las que la conciencia está presente y aquellas en que no lo está.

Por ejemplo, tanto William James como Jean Piaget se inclinan a considerar que la mentalidad es una característica de todas las cosas vivientes. James interpreta que la presencia de la mentalidad es, fundamentalmente, una cuestión del comportamiento inteligente o dirigido a un objetivo, que él opone al comportamiento meramente mecánico:

> *La búsqueda de fines futuros y la elección de los medios para alcanzarlos son la marca y el criterio de la presencia de la mentali-*

> *dad* en un fenómeno. Todos empleamos esta prueba para discriminar entre una actuación inteligente y una mecánica.[1]

De manera similar, aunque incluso con más énfasis, Piaget interpreta la diferencia entre lo mental y lo no mental a partir de la existencia de un propósito:

> No hay ningún tipo de frontera entre lo viviente y lo mental, o entre lo biológico y lo psicológico. [La psicología] no es la ciencia de la conciencia, sino del comportamiento en general [...] de la conducta. [La psicología comienza] cuando el organismo actúa en relación con situaciones externas y resuelve problemas.[2]

Las poderosas corrientes del pensamiento, entonces, se alejan de la suposición de que ser consciente es esencial para la mentalidad. Por supuesto, la ampliación de la mente propuesta por el psicoanálisis de modo que incluya los fenómenos inconscientes no requiere, por sí misma, que se les atribuya mentalidad a las criaturas que son completamente *incapaces* de tener conciencia. En cambio, las concepciones de James y de Piaget sí implican que la mentalidad es un rasgo característico de la vida de gran cantidad de criaturas –no sólo animales, sino también plantas–, que no gozan en absoluto de ninguna experiencia consciente.[3]

Ahora bien, ¿qué es esta *conciencia*, que se diferencia de la mentalidad y que, por lo general, suponemos que es privativa de los seres humanos y de los miembros de ciertas especies animales relativamente avanzadas? Anthony Kenny propone lo siguiente:

> Pienso que la conciencia [...] equivale a tener ciertos tipos de capacidad. Ser consciente es, por ejemplo, ver y oír. La cuestión

1 W. James, *The principles of psychology,* Cambridge, Mass., Harvard University Press, 1983, p. 21. [trad. esp.: *Principios de psicología*, México, Fondo de Cultura Económica, 1989].

2 J. C. Brinquier, *Conversations with Piaget,* Chicago, University of Chicago Press, 1980, pp. 3-4 [trad. esp.: *Conversaciones con Jean Piaget,* Barcelona, Gedisa, 1985].

3 Piaget mismo menciona el comportamiento de los girasoles como indicativo de mentalidad.

> de si alguien puede ver u oír se vincula con la de si puede discriminar entre ciertas cosas, y la cuestión de si puede discriminar entre ciertas cosas es algo que podemos probar tanto por medios simples y cotidianos como con complicados métodos experimentales.[4]

Kenny sugiere que ser consciente es ser capaz de discriminar. Pero ¿qué es discriminar? Parecería que discriminar entre dos cosas es, en el sentido fundamental, verse afectado de manera diferente por una que por la otra. Si mi estado continúa siendo exactamente el mismo sin que importe si determinada característica está presente o ausente en mi entorno, no discrimino entre la presencia y la ausencia de esa característica. Si mi estado cambia según esté presente o no esa característica, ése es un modo de discriminar entre su presencia y su ausencia. Discriminar sonidos, colores, niveles de temperatura y otras cosas similares sólo significa –en el sentido más general– responder a ellos de manera diferente.

Parece indiscutible que la posibilidad de discriminar es una característica central de la conciencia: ver implica necesariamente responder a las diferencias de color, oír, a las diferencias de sonido, y así sucesivamente. Esta caracterización, no obstante, no es en absoluto efectiva para dar cuenta de la idea que comúnmente tenemos de la conciencia. La manera habitual de identificar el estado de estar consciente es contrastándolo con la inconciencia, y una forma de estar inconsciente es estar dormido. Pero incluso dormidos, los animales responden a los estímulos visuales, auditivos, táctiles y a otros. De lo contrario, resultaría difícil despertarlos. Con seguridad, el rango de las respuestas que pueden dar cuando están durmiendo es más limitado que cuando están despiertos. Sin embargo, cuando duermen, no carecen por completo de la capacidad de discriminar y, por consiguiente, Kenny no puede considerar que en ese momento están completamente inconscientes.

Ahora bien, sería aceptable considerar que el sueño puede coexistir con cierto nivel de conciencia, inferior al del estado de vigilia, pero superior a cero. En la propuesta de Kenny, sin embargo, no

4 A. Kenny *et al.*, *The nature of mind*, Edimburgo, University Press, 1972, p. 43.

sólo los animales dormidos están conscientes, lo mismo sucede con todo lo demás en el mundo. Después de todo, no hay entidad que no sea susceptible de verse afectada en forma diferencial por algo. Si se interpreta que la noción de conciencia tiene el mero sentido general y primitivo que le atribuye Kenny, un trozo de metal es consciente de la temperatura ambiente en la medida en que se caliente o se enfríe, o se expanda y se contraiga, de acuerdo con los cambios de temperatura. La conciencia, así interpretada, es un estado respecto del cual el estado contrastante no es, claro está, la inconciencia, si entendemos por inconciencia lo que comúnmente le atribuimos a aquellos que están profundamente dormidos, anestesiados o en coma. En cambio, el estado con el cual contrasta la conciencia en este sentido es el aislamiento causal.

Por consiguiente, en el sentido ordinario de la palabra, la conciencia no puede ser exclusivamente una cuestión de discriminación, dado que (por así decirlo) en la oscuridad pueden producirse respuestas con capacidad de discriminar. Uno incluso podría evitar esta dificultad diciendo que la conciencia es la capacidad de discriminar en forma *consciente*, aunque ello no resultaría útil. En todo caso, quisiera considerar otra característica, diferente de la capacidad de discriminar, que es esencial para la conciencia ordinaria: la *reflexividad*. Ser consciente necesariamente implica no sólo diferenciar las respuestas a los estímulos, sino percibir esas respuestas. Cuando estoy despierto en un día caluroso, el calor aumenta la temperatura de mi piel; también aumenta la temperatura superficial de un trozo de metal. Tanto el metal como yo respondemos al calor y, en este sentido, ambos lo percibimos. Sin embargo, también percibo mi respuesta, mientras que el metal no lo hace. El aumento de la temperatura de mi piel es, en sí mismo, algo que discrimino, y esto es esencial para el modo de ser consciente que consiste en tener calor.

Por supuesto, el hecho de que una criatura responda a sus propias respuestas no implica que sea consciente. No hace falta decir que la segunda respuesta puede no ser más consciente que la primera. De esta manera, añadir la reflexividad a la discriminación no permite explicar cómo surge la conciencia ni en qué se diferencia de la inconciencia. No obstante, ser consciente en el sentido ordi-

nario de la palabra implica (a diferencia de la inconciencia) la reflexividad: necesariamente supone una percepción secundaria de una respuesta primaria. Una instancia de conciencia exclusivamente primaria e irreflexiva no constituiría una instancia de lo que comúnmente interpretamos por conciencia. Porque ¿cómo se podría ser consciente de algo sin percibir esta conciencia? Esto significaría vivir una experiencia sin percibir en absoluto que está ocurriendo. Éste sería, precisamente, un caso de experiencia inconsciente. Entonces, parece que ser consciente es lo mismo que ser autoconsciente. La conciencia *es* la autoconciencia.[5]

La afirmación de que la conciencia alerta es la autoconciencia no significa que la conciencia sea invariablemente dual en el sentido de que todas sus instancias involucran tanto una percepción primaria como otra instancia de conciencia que, de alguna manera, es distinta y separable de la primera, y que tiene a ésta como objeto. Ello sería el presagio de una proliferación intolerablemente infinita de instancias de conciencia. En cambio, la autoconciencia en cuestión es una especie de *reflexividad inmanente* en virtud de la cual cada instancia de ser consciente capta no sólo lo que percibe, sino también la percepción de ello. Se asemeja a una fuente de luz que, además de iluminar todas las otras cosas que están dentro de su alcance, también se torna visible.

II

Existe un problema desconcertante relacionado con la pregunta de *para qué* es la conciencia. Asimismo, es igualmente desconcertante que la función de la conciencia continúe siendo tan desconcer-

5 Cuando me refiero a la *autoconciencia*, no me refiero a la conciencia de un ser –un sujeto o un *yo*– ni a la conciencia de que hay percepción. Ambas requieren capacidades racionales más allá de lo que parecería ser necesario para que ocurra la conciencia misma. La reflexividad en cuestión es meramente la percepción que la conciencia tiene de sí misma. Oír un sonido conscientemente en vez de responder a él en forma inconsciente supone ser consciente de oírlo o ser consciente del sonido tal como se lo oye.

tante. Parece extraordinario que, a pesar de la omnipresencia y la familiaridad de la conciencia en nuestras vidas, no estemos seguros de cuál es el modo (si es que lo hay) en que nos es realmente indispensable.[6] Sea como fuere, la importancia de la *reflexividad* para aquellos en cuyas vidas ocurre es evidente. La sensibilidad de una criatura a su propia condición –ya sea mediante la introspección o la reflexividad inmanente de la conciencia alerta, o mediante una variedad menos deslumbrante de capacidad de respuesta secundaria– es esencial para el comportamiento tendiente a un fin determinado.

El metal no cambia con un fin determinado cuando se calienta; por otra parte, en ciertas condiciones, un girasol se vuelve hacia la luz. Tanto el metal como el girasol responden a lo que sucede a su alrededor. Cada uno de ellos es afectado por los estímulos ambientales y, por tanto, los discrimina. Sin embargo, el girasol, a diferencia del metal, hace discriminaciones de segundo orden además de las primarias. Esto contribuye fundamentalmente a su capacidad para el cambio con un fin determinado. El metal carece de esta capacidad, puesto que es insensible a sus propias respuestas, lo que significa que carece por completo de capacidad de respuesta o es indiferente a lo que le sucede. Una criatura con capacidad de respuesta secundaria está controlando su propia situación; en esa medida, la criatura está en condiciones de hacer algo acerca de su situación, o, al menos, se encuentra más cerca de estarlo.

6 Así, el fisiólogo laureado con el Premio Nobel, John Eccles dice: "Quisiera [preguntar], como neurofisiólogo, ¿por qué tenemos que ser conscientes? En principio, podemos explicar todas nuestras acciones de producción y recepción en términos de la actividad de los circuitos neuronales; y, por consiguiente, la conciencia parece absolutamente innecesaria. No creo en este cuento, por supuesto; pero a la vez no conozco una respuesta lógica para él. Al intentar responder la pregunta de por qué tenemos que ser conscientes, con seguridad no se puede afirmar que es evidente que la conciencia sea un requisito necesario para acciones tales como el razonamiento o la argumentación lógica, o incluso para la iniciativa y las actividades creativas", en J. Eccles (ed.), *Brain and consciuosness experience*, Nueva York, Springer-Verlag, 1964. Quizás, a pesar de la renuencia de Eccles a admitirlo, la instrospección de la vida humana es un absurdo ontológico, algo que se toma a sí mismo con enorme seriedad, pero que, en realidad, no tiene ninguna función importante.

De esta manera, la reflexividad tiene un sentido, como lo tiene la acción misma, debido al riesgo presente en la existencia. Le permite a una criatura, entre otras cosas, responder a la circunstancia de que sus intereses son afectados en forma adversa. Ello hace que la reflexividad sea una condición indispensable del comportamiento tendiente a evitar o a mejorar las circunstancias de este tipo, en las que hay un conflicto entre los intereses de una criatura y las fuerzas que los están haciendo peligrar o los están debilitando.

Asimismo, hay otro tipo de reflexividad o autoconciencia que, de manera similar, parece ser inteligible por ser fundamentalmente una respuesta al conflicto y al riesgo. Es una característica destacada de los seres humanos, que afecta nuestras vidas en forma profunda y de maneras innumerables, que nos preocupe lo que somos. Ello se relaciona estrechamente, como causa y como efecto, con nuestra inmensa preocupación acerca de lo que los demás piensan de nosotros. Estamos continuamente alerta ante el peligro de que pueda haber discrepancias entre lo que queremos ser (o lo que queremos que parezca que somos) y la impresión que les damos a los demás y a nosotros mismos.

En particular, nos preocupan nuestros propios motivos. Nos importa enormemente si los deseos por los que somos inducidos a actuar como lo hacemos nos motivan porque queremos que sean efectivos al hacerlo o si nos inducen independientemente de nosotros o, incluso, a pesar de nosotros. En los últimos casos, somos inducidos a actuar como lo hacemos sin querer ser motivados incondicionalmente como lo somos. Nuestro ánimo, en el mejor de los casos, está dividido, y es posible que ni siquiera participe de lo que estamos haciendo.

Esto significa, además, que hasta cierto punto somos pasivos respecto de la acción que realizamos. Debido a que no respaldamos o apoyamos en forma indudable nuestra propia motivación, se puede decir con propiedad que lo que queremos –es decir, el objeto de nuestro deseo motivador y el deseo mismo– no es, en sentido ordinario, algo que *realmente* queremos. Por tanto, si bien es posible que realicemos nuestra acción debido a la fuerza motivadora de nuestro propio deseo, también es verdad que estamos siendo inducidos a actuar por algo distinto de lo que en realidad

queremos. En ese caso, en cierta forma somos pasivos respecto de lo que nos induce, tal como lo somos cuando somos inducidos por una fuerza que no es por completo la nuestra.

Es posible que un ser humano sea a veces –e incluso siempre– indiferente a sus propios motivos; es decir, que no asuma una actitud evaluadora hacia los deseos que lo inclinan a actuar. De existir un conflicto entre esos deseos, no le importa cuál de ellos resulta el más efectivo. En otras palabras, el individuo no participa en el conflicto. Por tanto, el resultado del conflicto no le representa ni una victoria ni una derrota. Debido a que él no ejerce ninguna autoridad –mediante cuyo respaldo o aprobación algunos de sus deseos adquirirían una legitimidad particular o podrían tornarse especialmente constitutivos de su persona–, las acciones engendradas por el flujo y la confrontación de sus sentimientos y sus deseos son bastante inconscientes.

III

Ahora bien, ¿qué concepción de esta serie de fenómenos se ajusta a sus contornos de manera más auténtica y perspicua? En particular, prefiero un modelo que incluya niveles de reflexividad y de autoconciencia. Según este esquema, en el nivel más bajo hay deseos de primer orden de llevar a cabo una u otra acción. Cualquiera de estos deseos de primer orden que en realidad conduzca a la acción es designado, en virtud de esa efectividad, la voluntad del individuo a quien el deseo pertenece. Más aun, por lo general las personas tienen deseos de segundo orden relacionados con los deseos de primer orden que quieren, y tienen voliciones de segundo orden relacionadas con qué deseo de primer orden quieren que constituya su voluntad. Es posible que también haya deseos y voliciones de órdenes superiores.

Ello torna natural distinguir entre dos formas en que los aspectos volitivos de la vida de una persona pueden estar radicalmente divididos o ser incoherentes. En primer lugar, puede existir un conflicto entre el modo como alguien quiere ser motivado y el de-

seo por el cual es, de hecho, más poderosamente inducido. Un ejemplo de este tipo de conflicto interno lo brinda la situación de una persona que quiere abstenerse de fumar –es decir, que quiere que el deseo de abstenerse de fumar sea lo que, en efecto, motive su comportamiento–, pero cuyo deseo de fumar resulta ser tan fuerte que se convierte en su voluntad, a pesar del hecho de que la persona prefiera no satisfacerlo e, incluso, luche contra él. Aquí, hay una falta de coherencia o de armonía entre la volición de orden superior o la preferencia de la persona respecto de cuál de sus deseos quiere que sea más efectivo y el deseo de primer orden que, de hecho, es el más efectivo para inducirlo a la acción. Debido a que la persona preferiría no satisfacer el deseo que prevalece, el resultado del conflicto que hay dentro de ella es que no puede hacer lo que realmente quiere hacer. Su voluntad no está bajo su propio control. No es la voluntad que quiere, sino la que le impone una fuerza con la cual no se identifica y que, en ese sentido, es externa a ella.

Otro tipo de conflicto interno ocurre cuando no hay coherencia en el ámbito de las voliciones de orden superior de la persona. Esto no atañe a la relación entre las voliciones y la voluntad. No es una cuestión de fuerza volitiva, sino de si las preferencias de orden superior relacionadas con alguna cuestión volitiva son *incondicionales*. Tiene que ver con la posibilidad de que no haya una respuesta inequívoca a la pregunta de qué quiere en realidad la persona, aunque sus deseos formen una estructura jerárquica compleja y extensiva. Podría no haber una respuesta inequívoca, porque la persona es *ambivalente* con respecto al objeto que está más cerca de querer en realidad; en otras palabras, se debe a que la persona es atraída hacia ese objeto, pero también es alejada de él. O podría no haber una respuesta inequívoca porque las preferencias de la persona respecto de lo que quiere no están completamente integradas, de manera que hay cierta *incongruencia* o *conflicto* entre ellas, quizá aún no manifiestos.

Se podría decir que la incoherencia del primer tipo (del tipo que aqueja al fumador) está *entre* lo que la persona en realidad quiere y otros deseos –como el deseo de fumar, que, a pesar de ser rechazado, continúa predominando insoslayablemente–, que son *externos* al complejo volitivo con el que la persona se identifica y que

quiere que determine su comportamiento. El segundo tipo de incoherencia se encuentra *dentro* de este complejo volitivo. Por falta de incondicionalidad, la persona no sólo está en conflicto con las fuerzas "exteriores" a ella; más bien, ella misma está dividida.

Una ventaja de este modelo es que brinda una forma conveniente de explicar –como sucede en el caso del fumador renuente– cómo la pasividad o la autonomía debilitada pueden ser consecuencias de la fuerza de lo que, en un sentido básicamente literal, son los propios deseos del individuo. El modelo también se presta de maneras bastante obvias a la expresión y la explicación de una variedad de conceptos útiles relacionados con las características estructurales de la mente (por ejemplo, la debilidad de la voluntad, el ideal del yo, y otras). Sin embargo, la noción central del modelo de una jerarquía de deseos no parece ser completamente adecuada para su fin. La razón de esto es que parece imposible explicar, usando sólo los recursos de esta noción, de qué manera un individuo con deseos o voliciones de segundo orden puede ser menos inconsciente respecto de *ellos* que una criatura totalmente irreflexiva respecto de sus deseos de primer orden.[7]

Alguien hace lo que *realmente quiere* hacer sólo cuando actúa de acuerdo con una volición de orden superior pertinente. Sin embargo, esta condición podría no ser suficiente a menos que la persona *realmente quisiera* estar determinada por esa volición de orden superior en particular. Ahora bien, es muy claro que este requisito no puede satisfacerse simplemente introduciendo *otro* deseo o volición en el nivel superior siguiente. Ello produciría un movimiento recursivo que sería bastante arbitrario detener en algún punto particular. La dificultad afecta los dos tipos de incoherencia volitiva que he distinguido. Una caracterización de cualquiera de los tipos de incoherencia requiere interpretar que algunos de los deseos de una

7 La noción de reflexividad me parece mucho más importante y necesaria, cuando se tratan los fenómenos en cuestión, que la de jerarquía. Por otra parte, no es evidente que se pueda dar cuenta de la reflexividad sin recurrir a la noción de un orden jerárquico. Si bien expresar la vida volitiva en términos de una jerarquía de deseos parece algo artificial, las alternativas –como la que propone Gary Watson en "Free Agency", *Journal of Philosophy*, 1975– me parecen peores: más oscuras, no menos antojadizas y –sospecho– que al final también necesitan recurrir a la jerarquía.

persona se integran en ella de una manera en que los otros no lo hacen. No obstante, no resulta obvio cómo explicar la distinción entre los elementos volitivos que están integrados en una persona y aquellos que, en un sentido relevante, siguen siendo externos a ella.

El mero hecho de que un deseo ocupe un nivel más alto que otro en la jerarquía parece claramente insuficiente para dotarlo de mayor autoridad o de una legitimidad constitutiva. En otras palabras, la asignación de los deseos a diferentes niveles jerárquicos no brinda, en sí misma, una explicación de lo que para alguien significa *identificarse* con uno de sus deseos más que con otro. No deja en claro por qué debería ser adecuado interpretar que una persona *participa* en conflictos que tienen lugar en su interior entre voliciones de segundo orden y deseos de primer orden, y que, por tanto, es vulnerable a ser vencida por sus propios deseos, cuando, por otro lado, no se debe interpretar que un agente inconsciente es un participante genuino en los conflictos que ocurren en su interior entre deseos que son, todos, del primer orden, o que no tiene interés en sus resultados. Gary Watson expresó la cuestión en forma sucinta: "Debido a que las voliciones de segundo orden son, en sí mismas, simplemente deseos, añadirlas al contexto del conflicto significa aumentar la cantidad de contendientes; no significa asignarle un lugar especial a ninguno de los participantes de la disputa".[8] Parecería que el modelo jerárquico no puede, como tal, hacer frente a esta dificultad. Sólo nos permite señalar que hay un conflicto interno entre deseos de diferentes órdenes. Sin embargo, esto solo no es adecuado para determinar –en relación con ese conflicto– dónde se ubica la persona (si es que lo hace en algún lugar).[9]

Hace un tiempo, intenté abordar este problema en el siguiente pasaje:

> Cuando una persona se identifica *decididamente* con uno de sus deseos de primer orden, este compromiso "resuena" a través de

8 G. Watson, *op. cit.*, p. 128.

9 El problema de explicar la identificación no es, por supuesto, exclusivo del modelo jerárquico. Cualquier descripción de la estructura de la volición debe enfrentarse con él. Por tanto, no es una falla del modelo jerárquico que requiera una explicación de la identificación.

> la serie potencialmente infinita de órdenes superiores. [...] El hecho de que su volición de segundo orden de ser inducida por este deseo es decisiva significa que no hay lugar para preguntas acerca de la pertinencia de deseos o voliciones de órdenes superiores. [...] La firmeza del compromiso que ha contraído significa que ha decidido que no queda por formular ninguna otra pregunta acerca de su volición de segundo orden, en ningún orden superior.[10]

El problema de este pasaje es que las nociones a las que recurrí –es decir, *identificación*, *compromiso decisivo*, *resonar*– son terriblemente oscuras. Por consiguiente, el pasaje no permitía dilucidar justamente cómo la maniobra de evitar un regreso al infinito asumiendo un compromiso decisivo puede dejar de ser inaceptablemente arbitraria. Así, dice Watson:

> Queríamos saber qué es lo que evita la inconciencia respecto de las voliciones de orden superior que uno tiene. ¿Qué les da a estas voliciones una relación especial con "uno" mismo? No resulta útil responder que uno asume un "compromiso decisivo" cuando esto sólo significa que no se va a permitir un ascenso interminable a órdenes superiores. Esto *es* arbitrario.[11]

Ahora bien, de hecho, Watson está equivocado en este punto. Como intentaré explicar, asumir un compromiso decisivo no consiste meramente en una *negativa* arbitraria a permitir un ascenso interminable a órdenes superiores.

IV

Consideremos una situación de cierta manera análoga a la de una persona que duda de si identificarse con uno u otro de sus deseos,

10 "La libertad de la voluntad y el concepto de persona", capítulo 2 de este volumen.
11 G. Watson, *op. cit.*, p. 218.

pero bastante más directa: la situación de alguien que intenta solucionar un problema de aritmética. Después de hacer el cálculo, es posible que esta persona haga otro para verificar la respuesta. El segundo cálculo puede ser exactamente el mismo que el primero, o puede ser equivalente a él en el sentido de que sigue un procedimiento diferente del primero, pero debe arrojar el mismo resultado. De todas maneras, supongamos que el segundo cálculo confirma el primero. Es posible que ambos cálculos tengan errores, por lo que la persona puede volver a verificarlos. Esta secuencia de cálculos puede extenderse de manera indefinida. Más aun, no hay nada en la posición de ningún elemento de la secuencia en particular que le dé autoridad definitiva. Se puede cometer un error en cualquier instancia, y el mismo error puede repetirse cualquier cantidad de veces. Por tanto, ¿qué es lo que distingue a un cálculo con el que la persona puede dar por terminada la secuencia de manera razonable? ¿Cómo evita la persona ser irresponsable o arbitraria cuando en algún punto en particular pone fin a una secuencia que podría extender más?

Una manera en que podría terminar una secuencia de cálculos es que la persona que la realiza simplemente abandone la tarea y decida con negligencia que el resultado de su último cálculo le servirá como respuesta. Quizá pierda interés en el problema o quizá su atención se desvíe de continuar la indagación debido a alguna distracción irresistible. En casos como éstos, su comportamiento se asemeja al de un agente inconsciente: permite que cierto resultado aparezca como válido sin evaluar si es adecuado ni considerar la conveniencia de permitir que ésa sea su respuesta. No *elige* ni *aprueba* un resultado. Actúa como si le fuera totalmente indiferente la cuestión de si existe un respaldo apropiado para la aceptabilidad de su respuesta.

Por otra parte, una secuencia de cálculos podría terminar porque la persona que la está realizando *decide, por alguna razón*, adoptar cierto resultado. Puede ser que confíe en que ese resultado es correcto y, por tanto, crea que no tiene sentido continuar indagando. O quizá crea que, aunque haya alguna probabilidad de que el resultado no sea correcto, el costo que le representaría seguir indagando –en tiempo, en esfuerzo o en oportunidades perdidas– es

mayor que el valor que tiene para ella reducir la probabilidad de error. En cualquiera de los casos, puede haber una identificación "decisiva" de su parte. En un sentido que intentaré explicar, dicha identificación resuena a través de una secuencia ilimitada de posibles reconsideraciones posteriores de su decisión.

Supongamos que la persona confía en que sabe la respuesta correcta. Entonces, espera obtener esa respuesta cada vez que realiza con precisión un cálculo adecuado. En este aspecto, el futuro le es transparente, y su decisión de que cierta respuesta es correcta resuena sin cesar justamente en este sentido: le permite anticipar los resultados de una cantidad indefinida de otros cálculos posibles. Ahora bien, supongamos que no está totalmente segura de cuál es la respuesta correcta, pero que está convencida de que, sin embargo, sería más razonable de su parte adoptar cierta respuesta como propia. En ese caso, no puede esperar con total confianza que su respuesta sea confirmada por posteriores indagaciones; reconoce que un cálculo preciso podría arrojar un resultado diferente. Sin embargo, si ha asumido un compromiso genuinamente incondicional respecto de la opinión de que adoptar la respuesta es su alternativa más razonable, puede esperar que *esta* opinión sea confirmada indefinidamente por revisiones precisas de ella.

El hecho de que un compromiso resuene indefinidamente *significa,* simplemente, que el compromiso es *decisivo.* Un compromiso es decisivo si y sólo si se asume sin reservas, y asumir un compromiso sin reservas significa que la persona que lo asume lo hace porque cree que ninguna otra indagación precisa le exigiría que cambiara de parecer. Por consiguiente, no tiene sentido continuar con la indagación. En esto consiste, precisamente, el efecto de la resonancia.[12]

Ahora bien, lo que conduce a las personas a formar deseos de órdenes superiores es similar a lo que las conduce a repasar sus

12 Aquí, concuerdo con el planteamiento sobre la relación entre la resonancia y el compromiso decisivo que hace Jon Elster en *Ulysses and the sirens: Studies in rationality and irrationality,* Cambridge, Cambridge University Press, 1979, p. 111, n. 135 [trad. esp.: *Ulises y las sirenas. Estudios sobre racionalidad e irracionalidad,* México, Fondo de Cultura Económica, 1989]. Mi propio tratamiento de estos asuntos le debe mucho a la exposición de Descartes sobre la percepción clara y distinta.

cálculos aritméticos. Alguien verifica sus cálculos porque piensa que puede haberlos hecho mal. Puede ocurrir que exista un conflicto entre la respuesta que obtuvo y una respuesta diferente que, por una razón u otra, cree que puede ser correcta; o quizá simplemente sospeche, de un modo más general, que puede haber cometido algún tipo de error. De manera similar, una persona puede ser llevada a reflexionar sobre sus propios deseos, ya sea porque discrepan unos con otros, ya porque una falta de confianza más general la induce a considerar si debe sentirse satisfecha con sus motivos tal como son.

Tanto en el caso de los deseos como en el caso de los cálculos, una persona puede dar por terminada sin arbitrariedad una secuencia potencialmente infinita de evaluaciones cuando considera que no hay ningún conflicto inquietante, ya sea entre los resultados que obtuvo, ya entre un resultado obtenido y otro que, con razón, podría esperar obtener si la secuencia continuara. Dar por terminada la secuencia en ese punto –el punto en que no hay ni conflicto ni duda– no es arbitrario. Ello se debe a que la única razón para continuar la secuencia sería hacer frente a un conflicto real o a un conflicto potencial. Dado que la persona no tiene esta razón para continuar, no es arbitrario que se detenga.

Puede sugerirse que sigue habiendo un elemento de arbitrariedad aquí, en la consideración de que no se puede encontrar ningún conflicto pertinente: después de todo, esta consideración también es susceptible de error, y sería posible reevaluarla una y otra vez sin que ninguna de las reevaluaciones fuera intrínsecamente definitiva o final. Sin embargo, por más sólido que sea este argumento, no supone una deficiencia específica del principio de que una persona tiene motivos para dar por terminada una secuencia de cálculos o de reflexiones cuando no se encuentra con conflictos que deban evitarse o resolverse. El argumento es bastante general. *Siempre* es posible, al formular cualquier principio, hacer el juicio erróneo o sin fundamento de que se han satisfecho las condiciones para aplicar dicho principio correctamente. No debería hacer falta decir que ningún criterio ni ninguna norma pueden garantizar que el principio será esgrimido con precisión y sin arbitrariedad.

V

El significado etimológico del verbo *decidir* es "cortar". Esto resulta acertado, ya que es habitual (aunque no necesario, y ni siquiera muy frecuente) que una decisión ponga fin a una secuencia de deseos o de preferencias pertenecientes a órdenes cada vez más altos. Cuando se toma la decisión sin reservas, el compromiso que trae como consecuencia es decisivo. Entonces la persona ya no se aleja del deseo con el cual se ha comprometido. Ya no queda sin definir si el objeto de ese deseo –en otras palabras, lo que la persona quiere– es lo que realmente quiere: la decisión determina lo que la persona en realidad quiere apropiándose plenamente del deseo sobre el cual basa su decisión. En esta medida, la persona, al tomar una decisión por la cual se identifica con un deseo, *se constituye.* El deseo pertinente ya no es, de ninguna manera, externo a ella. No es un deseo que meramente "tiene" como sujeto en cuya historia casualmente sucede, tal como una persona puede "tener" un espasmo involuntario que casualmente se produce en la historia de su cuerpo. Llega a ser un deseo que se incorpora a ella en virtud del hecho de que lo tiene *por su propia voluntad.*

Esto no significa que el deseo se origina mediante el ejercicio de la voluntad; el deseo bien puede existir antes de que se tome una decisión respecto de él. No obstante, incluso si la persona no es responsable del hecho de que el deseo *ocurra*, en un sentido importante asume la responsabilidad de tener el deseo –el hecho de que el deseo sea suyo en el sentido más amplio, que éste constituya lo que la persona realmente quiere– cuando se identifica con él. Mediante su acción de decidir, es responsable del hecho de que el deseo se ha vuelto suyo de una manera que, claramente, antes no lo era.

Existen dos tipos de conflictos bastante diferentes entre los deseos. En los conflictos de uno de los tipos, los deseos compiten por la prioridad o la posición en un orden de preferencias; la cuestión es qué deseo satisfacer *en primer lugar*. En los conflictos del otro tipo, la cuestión es si debería asignársele a un deseo *algún* lugar en el orden de preferencia; es decir, si se lo debe aprobar como candidato legítimo para ser satisfecho o si se lo debe rechazar por no tener derecho a ningún tipo de prioridad. Cuando se resuelve un

conflicto del primer tipo, los deseos en competencia son *integrados* en una única jerarquía, dentro de la cual cada uno ocupa una posición específica. Para resolver un conflicto del segundo tipo, se requiere una *separación* radical de los deseos en competencia, a uno de los cuales no sólo se le asigna una posición relativamente menos favorecida, sino que se lo excluye por completo como a alguien que está fuera de la ley. Estos actos de jerarquización y de rechazo –integración y separación– son los que crean una identidad a partir de las materias primas de la vida interior. Definen las restricciones y los límites intrapsíquicos respecto de los cuales la autonomía de una persona puede verse amenazada incluso por sus propios deseos.[13]

Aristóteles sostenía que el comportamiento es voluntario sólo cuando su principio motor se encuentra dentro del agente. Esto no puede ser correcto si se interpreta el término *dentro* en su sentido literal: los movimientos de un ataque epiléptico no son voluntarios, pero su principio motor o causa es, desde un punto de vista espacial, interna al agente. La ubicación de un principio motor respecto del cuerpo del agente es claramente menos relevante que su "ubicación" respecto de la volición de éste. Lo que importa, incluso respecto de un principio motor que opera como elemento de su vida psíquica, es si el agente se ha constituido para incluirlo o no. Por un lado, el principio puede ser interno, en el sentido pertinente a la cuestión de si el comportamiento al cual conduce es voluntario, en virtud del hecho de que la persona se ha unido a lo que la induce mediante un compromiso a través del cual asume la responsabilidad de ello. Por otro lado, el principio motor de su comportamiento puede continuar siendo externo a la persona en el sentido pertinente, porque no lo ha convertido en parte de ella.

Esto sugiere otro aspecto en el que la teoría de Aristóteles es insatisfactoria. Él sostiene que una persona puede ser responsable de su propio carácter por haber tomado (o por no haber tomado)

13 Las condiciones determinantes pertinentes en este caso son exclusivamente arreglos *estructurales*. Lo menciono, aunque no profundizaré en el tema, porque tiene relación con la conocida cuestión de si las consideraciones *históricas* –en especial los relatos causales– tienen alguna relevancia esencial para las preguntas acerca de si las acciones de una persona son autónomas.

medidas que afectan sus disposiciones habituales. En otras palabras, según Aristóteles una persona adquiere la responsabilidad de su propio carácter cuando actúa de una manera que resulta causalmente instrumental para provocar que tenga el conjunto particular de disposiciones que constituyen su carácter. Pienso que el tratamiento que Aristóteles hace de este tema está significativamente desenfocado debido a su preocupación por los orígenes causales y la responsabilidad causal. La responsabilidad fundamental de un agente respecto de su propio carácter no es cuestión de si éste *tiene* cierta disposición para sentir y para comportarse de diversas maneras por efecto de sus propias acciones. Ello sólo concierne a la pregunta de si la persona es responsable de tener estas *características*. La pregunta de si la persona es responsable de su propio *carácter* tiene que ver con la pregunta acerca de si ha *asumido la responsabilidad de* sus características. Atañe al problema de si las disposiciones en cuestión, independientemente de si su *existencia* se debe o no a la propia iniciativa y a la agencia causal de la persona, son características con las cuales ella se identifica y que, por tanto, incorpora en sí por voluntad propia como constitutivas de lo que es.

Cuando alguien se identifica con uno de sus propios deseos y no con otro, el resultado no es necesariamente la eliminación del conflicto entre esos deseos, ni siquiera la reducción de su seriedad, sino la alteración de su naturaleza. Supongamos que una persona con dos deseos conflictivos se identifica con uno y no con el otro. Esto *podría* causar que el otro –el deseo con el que la persona no se identifica– se tornara sustancialmente más débil de lo que era o que desapareciera por completo. Sin embargo, no es necesario que esto ocurra. Es muy posible que el conflicto entre ambos deseos continúe tan virulento como antes. El compromiso de la persona con uno de los deseos no elimina el conflicto entre éste y el otro. Elimina el conflicto *dentro de la persona* respecto de cuál de estos deseos prefiere que sea su motivación. El conflicto entre los *deseos* es transformado, de esta manera, en un conflicto entre *uno* de ellos y la *persona* que se ha identificado con su rival. Esa persona ya no siente incertidumbre acerca de qué partido tomar en el conflicto entre los dos deseos, y la persistencia de este conflicto no tiene por

qué subvertir ni reducir la incondicionalidad de su compromiso para con el deseo con el que se identifica.

VI

Debido a que tomando una decisión es el modo en que una persona se identifica más conspicuamente con algún elemento de su vida psíquica, el tomar decisiones desempeña una función importante en la formación y el sostén del yo. Resulta difícil expresar con claridad en qué consiste el acto de decidir: dejar totalmente en claro qué hacemos cuando lo llevamos a cabo. Sin embargo, si bien la naturaleza de la decisión es irritantemente esquiva, al menos es evidente que tomar una decisión es algo que nos hacemos *a nosotros mismos.* En este aspecto, difiere en forma fundamental del hecho de elegir, cuyo objeto inmediato no es el que elige, sino lo que elige. Esta diferencia entre la decisión y la elección explica el hecho de que decidir hacer cierta elección no sea lo mismo que hacerla realmente (después de todo, es posible que el momento o la ocasión de hacerla no haya llegado aún), mientras que decidir tomar una decisión en particular (es decir, decidir decidir las cosas de cierta manera) no puede distinguirse de tomar la decisión misma.

En algunos idiomas, la reflexividad de decidir –el hecho de que se trata de una acción que uno se hace a sí mismo– se indica en la forma del verbo pertinente. Así, en francés el verbo es *se décider.* En inglés, existe la frase *to make up one's mind,* uno de los sinónimos de *to decide,* en la que hay una representación explícita del carácter reflexivo de la decisión.* Ahora bien, ¿qué debemos hacer con la metáfora algo proteica que evoca esta frase? ¿Acaso *decidirse* –expresado como *to make up one's mind*– se puede equiparar a *inventar una historia* (*to make up a story*) o a *hacer la cama* (*to make up a bed*)? ¿Es como *maquillarse* (*to make up one's face*) o como *hacer una lista de tareas* (*to make up a list of things to do*)? ¿O es como *hacer las paces* (*to make up after a quarrel*)? ¿Cuál es la diferencia, en

* En español, un equivalente reflexivo sería la expresión *decidirse.* [N. de T.]

estos diversos ejemplos, entre lo que se hace y lo que no? ¿Y cuál de estas diferencias se corresponde más con la diferencia entre una mente que se ha decidido y una que no lo ha hecho?

El uso de maquillaje se refiere a un contraste entre cómo se ve una persona naturalmente y cómo la persona puede ingeniárselas para verse. Un contraste similar está implícito en la idea de inventar una historia, que se asemeja a maquillarse en el sentido de que el resultado es, en ambos casos, algo artificial o ficticio; simplemente no muestra las cosas como son en realidad. Una diferencia entre maquillarse e inventar una historia es, por supuesto, que la cara existe antes de maquillarla; es algo que se maquilla. Esto no es fácilmente comparable con el caso de una historia, que no se transforma cuando se la inventa, sino que llega a existir sólo cuando se la idea. En este aspecto, maquillarse se asemeja más a hacer la cama. Respecto de hacer una lista, claramente no tiene nada que ver con lo ficticio o lo inventado, sino con establecer ciertas relaciones entre los elementos de la lista o de registrar las relaciones que ya existen entre ellos.

Lo que parece básicamente común a todas las ocurrencias mencionadas no es el contraste entre la ficción y la realidad o entre lo natural y lo artificial, sino el tema de crear un orden. A mi criterio, bajo esta luz, la situación más parecida a la de decidirse a hacer algo es, quizá sorprendentemente, aquella en la que dos personas hacen las paces. Las personas que hacen eso luego de una pelea pasan de una situación de conflicto y hostilidad a una relación más armoniosa y ordenada. Por supuesto, las personas no siempre hacen las paces cuando la pelea finaliza; a veces, su hostilidad continúa incluso después de que el conflicto que fue su causa original se resuelve. Más aun, las personas que han estado peleando pueden restaurar la armonía entre ellas aunque su desacuerdo continúe. Hacer las paces es enmendar una relación arruinada por el conflicto, y no tiene nada que ver –directa o necesariamente– con la cuestión de si el conflicto ha terminado o no.

Interpretada sobre la base de esta analogía, tomar una decisión parece diferir de las actividades autorreparadoras del cuerpo a las que, en algunos aspectos, se parece. Cuando el cuerpo se cura a sí mismo, *elimina* conflictos en los que un proceso físico (digamos

una infección) interfiere con otros y socava la homeostasis o equilibrio en el que consiste la salud. Una persona que toma una decisión también busca, de esa manera, superar o reemplazar una situación de división interna y tornarse un todo integrado. Sin embargo, puede lograr esto sin eliminar, en realidad, los deseos que entran en conflicto con aquellos sobre los cuales ha tomado una decisión, siempre que se disocie de ellos.

Es posible que una persona no logre integrarse cuando toma una decisión, por supuesto, dado que el conflicto o la indecisión a la que hace frente puede continuar a pesar de su decisión. Todo lo que una decisión hace es crear una intención; no garantiza que la intención se lleve a cabo. Esto no se debe simplemente a que la persona siempre puede cambiar de parecer. Más allá de la inconstancia de ese tipo, puede ser que las energías tendientes hacia una acción que es contradictoria con la intención permanezcan indómitas y unidas, sin importar cuán terminantemente la persona piense que se ha decidido. El conflicto que se suponía que la decisión iba a sustituir puede continuar, a pesar de la convicción de la persona de que lo ha resuelto. En ese caso, la decisión, independientemente de cuán consciente y sincera parezca, no es incondicional: sea consciente de ello o no, la persona tiene otras intenciones, intenciones incompatibles con la que estableció la decisión y con la cual también está comprometida. Esto puede tornarse evidente a la hora de la verdad, y la persona está ostensiblemente imposibilitada por la intención respecto de la cual pensaba que había tomado una decisión.

VII

No obstante, ¿por qué nos interesa decidirnos? Parecería que el sentido de decidir es asegurar la realización de una acción que, de lo contrario, no se llevaría a cabo. Supongamos que me decido a mostrar mi enojo más abiertamente la próxima vez que un funcionario arrogante me insulte en forma gratuita. Se podría interpretar que, de esta manera, se establece una relación que antes no existía entre el comportamiento insultante de cierto tipo y la clase

de respuestas respecto de la cual acabo de tomar una decisión: una relación tal que si se produce la provocación, tendrá una respuesta como consecuencia. De hecho, sin embargo, las personas a menudo deciden hacer cosas que –lo sepan o no– harían de todas maneras. La relación entre la provocación y la respuesta, que la decisión parecería establecer, puede existir de antemano: yo habría demostrado mi enojo abiertamente incluso si, con antelación, no hubiera establecido la intención de hacerlo. El sentido de decidirse no es, por tanto, asegurar una acción determinada.

Tampoco es asegurar que uno actuará correctamente. Ésa es la función de la deliberación, que está destinada a aumentar la probabilidad de que las decisiones sean correctas. Hobbes sugiere que la palabra *deliberación* connota una actividad en la que se pierde la libertad.[14] Es, después de todo, *de-liberación*. Esto puede parecer paradójico, dado que habitualmente consideramos que la deliberación está vinculada de manera paradigmática con el ejercicio de la autonomía. La dificultad desaparece cuando reconocemos que la libertad con la que la deliberación interfiere no es la del agente autónomo, sino la de alguien que sigue los impulsos a ciegas, en otras palabras, el agente inconsciente. Una persona que delibera acerca de qué hacer está buscando una alternativa a "hacer lo que surge en forma natural". Su objetivo es reemplazar la libertad del comportamiento impulsivo y anárquico por la autonomía que significa estar bajo su propio control.

Una de las cosas que logra una decisión deliberada, cuando crea una intención, es establecer una restricción por la cual se guiarán las otras preferencias y decisiones. Una persona que decide qué creer se forma un criterio para otras creencias; es decir, éstas deben ser coherentes con la creencia respecto de la cual ha tomado una decisión. Y una persona que toma una decisión respecto de qué hacer adopta, de manera similar, una regla para coordinar sus actividades con el fin de facilitar la posible instrumentación de la de-

14 T. Hobbes, *Leviatán*, parte I, cap. 6: "Y esto se llama *deliberación*, porque implica poner término a la libertad que tenemos de hacer u omitir, de acuerdo con nuestro propio apetito o aversión" [trad. esp.: México, Fondo de Cultura Económica, 1980 (2ª ed.), p. 48].

cisión que ha tomado. Se podría decir, entonces, que una función de la decisión es integrar a la persona desde una perspectiva tanto dinámica como estática. Dinámica, en la medida en que asegura –en la forma que acabo de mencionar– coherencia y unidad de propósito en el tiempo; estática, en la medida en que establece –como mencioné antes– una estructura reflexiva o jerárquica por la cual se puede constituir, en parte, la identidad de una persona.

En ambos aspectos, la intención es, al menos en parte, resolver el conflicto o evitarlo. Esto no se logra eliminando uno o más de los elementos en conflicto de manera que los que queden estén en armonía, sino respaldando o identificando ciertos elementos que, entonces, tienen autoridad para el yo. Por supuesto, esta autoridad puede ser resistida e, incluso, vencida por fuerzas ilegales: deseos o motivos por los que la persona no quiere ser inducida efectivamente, pero que no se pueden evitar porque son demasiado fuertes e insistentes. También puede ser que haya un conflicto dentro de la autoridad misma: que la persona se haya identificado de manera incoherente. Ésta es la cuestión de la *incondicionalidad.*

La incondicionalidad, tal como estoy empleando la palabra, no consiste en un sentimiento de entusiasmo o de certidumbre frente a un compromiso. Tampoco es probable que resulte inmediatamente evidente la cuestión de si una decisión que una persona tiene intenciones de que sea incondicional lo es en realidad. No nos conocemos lo suficientemente bien para estar seguros de que nuestra intención de que nada interfiera con la decisión que tomamos sea la que querremos que se mantenga cuando –quizás al darnos cuenta de que no hay vuelta atrás– lleguemos a comprender más íntegramente qué se requerirá que hagamos o dejemos de hacer para cumplirla.

Cuando toma una decisión, una persona establece preferencias respecto de la resolución de conflictos entre sus deseos o sus creencias. Alguien que toma una decisión de ese modo lleva a cabo una acción, pero no se trata de un simple acto que sólo instrumente un deseo de primer orden. En esencia, supone la reflexividad, que incluye deseos y voliciones de un orden superior. Así, las criaturas que son incapaces de esta reflexividad volitiva necesariamente carecen de la capacidad de decidirse. Pueden desear, pensar y actuar,

pero no pueden decidir. En la medida en que interpretemos que tomar decisiones es la función característica de la facultad de la volición, debemos considerar que dichas criaturas carecen de esta facultad.

En "La libertad de la voluntad y el concepto de persona" (capítulo 2 de este volumen), afirmé que ser inconsciente no impide la deliberación. Entonces pensaba que aunque una criatura pudiera ser inconsciente respecto de los objetivos, podría, de todos modos, realizar cálculos o razonar sobre cuestiones técnicas acerca de cómo obtener lo que desea a su manera inconsciente. Sin embargo, razonar supone tomar decisiones acerca de qué pensar, que no parecen menos incompatibles con la inconciencia extrema que decidir lo que uno quiere hacer. Tomar una decisión (*making a decision*) parece tener un significado diferente de darse cuenta de cómo implementarla, pero no está claro que esta última actividad pueda lograrse sin decidirse (*make up one's mind*), de un modo estructuralmente muy similar a la de la primera.

Estamos acostumbrados a pensar que nuestra especie se distingue, en particular, por la facultad de la razón. Tendemos a suponer que la volición o la voluntad es una facultad más primitiva o rudimentaria, que compartimos con criaturas de menor complejidad psíquica. Sin embargo, ello parece dudoso no sólo por la reflexividad que la volición misma requiere, sino también en la medida en que el razonamiento requiere decidirse. Ello se debe a que, hasta ese punto, el uso deliberado de la razón necesariamente tiene una estructura jerárquica que requiere elementos de orden superior de los que no dispone un agente inconsciente genuino. En este aspecto, entonces, la razón depende de la voluntad.

13
La racionalidad y lo impensable

I

Respecto de acciones de cualquier tipo, se pueden concebir circunstancias en las que una acción precisamente de ese tipo tenga mayor utilidad que cualquier otra alternativa. Esto significa que si el utilitarismo es correcto, cualquier cosa puede tornarse, en algún punto, moralmente imperativa. Hay personas que sostienen lo mismo acerca del ateísmo. Si Dios no existe, dicen, cualquier cosa vale. Una persona puede hacer o ser lo que le guste.

Estas observaciones sobre el ateísmo y el utilitarismo no son iguales, pero están estrechamente relacionadas. Sugieren que las doctrinas a las que atañen tornan imposible creer que existan límites morales absolutos. Se supone que el corolario del ateísmo es que nada está prohibido: si no hay Dios, *todo está permitido*. De la misma manera, se presume que el utilitarismo implica que *cualquier cosa puede ser necesaria*. Si suponemos que estas descripciones de ambas doctrinas son correctas, quienes no adhieren a ninguna de ellas no reconocen ninguna restricción moral incondicional. En otras palabras, los utilitaristas y los ateos concuerdan en que *no se puede descartar nada de antemano*.

Estamos acostumbrados a dar por sentado que la expansión de nuestra libertad nos enriquece. Lo hace, sin embargo, sólo hasta cierto punto. Si las restricciones a las elecciones que puede hacer una persona se flexibilizan demasiado, es posible que ésta se sienta desorientada e insegura acerca de qué y cómo elegir. Una gran proliferación de opciones puede debilitar la comprensión que tiene de

su propia identidad. Cuando se enfrenta a la tarea de evaluar y de clasificar una gran cantidad de alternativas, su apreciación previa de cuáles son sus intereses y sus prioridades puede tornarse menos decisiva. Puede debilitarse la confianza que tiene en sus propias preferencias, que había desarrollado cuando las posibilidades que tenía eran menos y más conocidas. Es decir, se puede perturbar su claridad y su confianza respecto de quién es.

Ahora bien, supongamos que el campo de alternativas entre las que una persona puede elegir no sólo se expande; supongamos que sus límites son borrados por completo. En otras palabras, supongamos que, ahora, la persona tiene a su disposición y puede elegir cualquier curso de acción, incluso aquellos que afectarían sus propias preferencias. Debido a que, en ese caso, ella puede incluso alterar su propia voluntad, al parecer tiene que enfrentar las elecciones que debe hacer sin ningún carácter volitivo específico que sea definitivamente suyo. La voluntad de la persona, suponemos, es la que ella elige que sea. Por consiguiente, no es nada hasta que ella decide qué voluntad elegir.

Pero, entonces, ¿cómo ha de hacer una elección? ¿Qué preferencias y prioridades la guiarán al elegir, cuando sus propias preferencias y prioridades se encuentran entre las cosas mismas que debe elegir? Parecería que le queda tan poca sustancia volitiva que no puede considerarse que ninguna elección que haga surja de su naturaleza genuina. Respecto de una persona cuya voluntad no tiene un carácter determinado fijo, parece que la noción de autonomía o de autodirección no puede encontrar dónde aferrarse. Una persona como ésta está tan vacía de tendencias y de limitaciones identificables que le resultará imposible deliberar o tomar decisiones a conciencia. Es posible que siga siendo capaz de hacer cierta hueca apariencia de elección. Sin embargo, si así fuera, sería sólo en virtud de un vestigio de susceptibilidad a incipientes espasmos volitivos. Y los movimientos de ese tipo de su voluntad son intrínsecamente demasiado arbitrarios para estar por completo vacíos de un auténtico significado personal.[1]

1 Durante algún tiempo, muchos de aquellos que han sido admirados ampliamente por estar entre los pensadores más iluminados y humanos han hecho caso omiso

Desde luego, nadie alega que los partidarios del utilitarismo también estén predestinados a ese paralizante vacío volitivo. Lo que *sí* se sostiene a veces es que los utilitaristas no pueden asumir compromisos morales sólidos. Se afirma que son incapaces de sostener concepciones significativas de integridad personal. Se aduce que ello se debe a que un utilitarista debe querer alterar sus principios, sus prioridades y todos los otros elementos que conforman su carácter así como modificar su comportamiento cuando hacerlo tuviera más utilidad que hacer otra cosa. Como lo expresa Rawls:

> Los miembros de una sociedad utilitarista bien ordenada [...] no tienen un concepto determinado del bien con el cual estén comprometidos, sino que consideran que los diversos deseos y capacidades del yo son características que deben ajustarse en la búsqueda [de la mayor utilidad].[2]

Ello parece implicar que un utilitarista no puede comprometerse realmente a mantener ningún tipo específico ni ningún conjunto particular de ideales morales. Por consiguiente, parece que no puede formar un concepto estable de su propia identidad moral.

Para el utilitarismo, el único bien racional es el bienestar. Ahora bien, el grado de bienestar de una persona es una función no sólo de sus circunstancias externas, sino también de sus características personales. Por tanto, es racional que un utilitarista modifique cualquiera de sus características personales, incluyendo sus compromisos con valores específicos, siempre que ello incremente el bienestar. Según Rawls, esto origina una concepción del yo o de la persona que no le asigna carácter inherente:

---

del incipiente conflicto o tensión entre la libertad y la identidad. Han demostrado una tendencia constante *tanto* a instar a la expansión general de la oportunidad y la elección *como* a otorgar un gran valor a la individualidad. Por tanto, no debería sorprender que las cosas no hayan marchado bien.

2 J. Rawls, "Social unity and primary goods", en Amartya Sen y Bernard Williams (eds.), *Utilitarianism and beyond*, Cambridge, Cambridge University Press, 1982, p. 180.

> [El utilitarismo] define a las personas como lo que podemos denominar *personas desnudas*. Dichas personas están dispuestas a considerar cualquier convicción y objetivos nuevos, e incluso a abandonar aficiones y lealtades, cuando hacerlo les promete una vida con mayor satisfacción o bienestar generales. [...] La noción de persona desnuda [...] representa la disolución del concepto de la persona como alguien cuya vida expresa el carácter y la devoción a fines últimos específicos y a valores adoptados (o afirmados) que definen los puntos de vista distintivos asociados a conceptos diferentes (e inconmensurables) del bien.[3]

El problema con el utilitarismo, según esta opinión, es que exige demasiada flexibilidad. En su caracterización de las personas, que es extraña y perturbadoramente impersonal, no parece haber lugar para una noción como la integridad moral. Ello se debe a que para un utilitarista, ningún aspecto de su naturaleza volitiva puede estar, en rigor, exento de ser alterado en forma deliberada. Por el contrario, debe estar preparado para modificarse a sí mismo de cualquier manera que tenga como resultado un incremento de su bienestar. Por tanto, no puede comprometerse a respetar como inviolable ningún límite que pueda servir a la vez para arraigar su juicio y para especificar los requerimientos de su integridad.

Es posible, entonces, argumentar que en su intento de decidir qué hacer o qué ser los ateos y los utilitaristas tienen muy poco a qué recurrir en cuanto a valores personales firmemente arraigados u otras guías estables para la elección. Carecen de lo que Rawls llama *estructura moral previa*.[4] Sus doctrinas les prohíben pensar (al menos, eso es lo que se argumenta) que tienen límites volitivos fijos. En ambos casos, un exceso de libertad causa una disminución o, incluso, una disolución de la realidad del yo.

3 J. Rawls, "Social unity...", *op. cit.*, pp. 180-181. La noción de *persona desnuda* se parece a la antigua idea teológica del alma: algo que es idéntico en todos los hombres, que no es afectado por la experiencia, que no podemos identificar con nada en nosotros mismos de lo que seamos conscientes, pero que, sin embargo, de cierta manera le otorga a cada uno de nosotros su naturaleza esencial y su valor último.

4 *Ibid.*, p. 182.

II

No obstante, en mi opinión Rawls se equivoca al sostener que los utilitaristas no pueden tener "una concepción definida del bien con el que están comprometidos". El argumento que esgrime es que, debido a que un utilitarista debe incluir sus propios valores personales entre "los deseos y las capacidades del yo [...] que deberán adaptarse" siempre que ello incremente el bienestar, nunca puede comprometerse en forma total o incondicional con esos valores. Sin embargo, esto parece pasar por alto un punto importante. Aunque en principio alguien esté dispuesto a adaptar sus valores siempre que ello incremente el bienestar, puede no estar, sin embargo, legítima y completamente convencido de que, en realidad, hay ciertos tipos de adaptaciones que nunca le permitirán incrementarlo. En otras palabras, es posible que una persona tenga buenas razones para confiar en que nunca se producirán las circunstancias que le exigirían hacer ciertas adaptaciones en sus propios deseos y capacidades. El hecho de que se puedan concebir circunstancias de cierto tipo no implica, por cierto, que nadie tenga pruebas suficientes para justificar la afirmación de que no se producirán.

No hay razones por las cuales un utilitarista que tiene un conjunto específico de valores personales no asuma un compromiso para con ellos tan incondicional como su esperanza de que nunca se enfrentará a circunstancias en las que mantener esos valores le exigiría sacrificar el bienestar. Si su esperanza al respecto es total, su compromiso tampoco necesita ser equívoco ni restringido. Sin poner en riesgo su integridad de ninguna manera, alguien que está comprometido con un principio de conducta puede reconocer que en determinadas circunstancias podría ser razonable violar el principio y que, de hecho, él lo violaría. Ello se debe a que si está totalmente convencido de que ninguna de esas circunstancias tendrá lugar, también estará convencido de que su principio nunca le exigirá que realice acciones que su compromiso prohíbe. De allí que no haya necesidad de que él considere que las acciones de ese tipo tienen algún lugar en su repertorio de conducta.

Por cierto, puede estar equivocado. En ese caso, contrariamente a sus expectativas, en algún momento se enfrentará con una situa-

ción en la que la única forma de mantener su adhesión al utilitarismo sea violando su compromiso para con otros valores. Sin embargo, esto no significa que no pueda asumir, en forma razonable, compromisos inequívocos tanto con el utilitarismo como con sus otros valores. La posibilidad de que las expectativas que uno tiene sean erróneas sólo significa que se corre un riesgo al basar en ellas un compromiso incondicional. Ello no implica que asumir el riesgo sea imposible o injustificado. Por tanto, el utilitarismo no impide que los compromisos personales se abstengan de tipos de conducta específicos (o que se involucren en ellos). Aunque en ciertas condiciones las conductas de esos tipos podrían maximizar (o no) la utilidad y, por tanto, serían obligatorias (o estarían prohibidas), puede ser *incuestionable* que esas condiciones nunca ocurrirán.

Se puede establecer un paralelo con el ateísmo. Incluso si el ateo cree que todo es admisible, quizá haya ciertas cosas que, simplemente, no logra hacer. Esas cosas pueden descartarse como posibles cursos de acción para él. Aunque podría ser capaz de realizarlas si eligiera hacerlo, y a pesar de que sabe que realizarlas está *permitido*, ellas no se encuentran entre sus opciones genuinas. No tienen ningún lugar en su repertorio de conducta porque, podríamos decir, para él son *impensables.*

III

¿Qué significa decir que una acción es, para cierta persona, impensable? ¿Qué tipo de incapacidad o de impedimento está en juego cuando la explicación del hecho de que alguien no haya realizado una acción es que no puede lograr llevarla a cabo? A veces, las personas son incapaces de hacer cosas porque las circunstancias no son las adecuadas o porque carecen del poder o de la habilidad necesarios. Nadie puede abandonar una habitación cuyas salidas están firmemente bloqueadas o hacerse invisible o tocar el piano cuando no sabe tocar o cuando no tiene el instrumento. Sin embargo, a veces una persona puede ser incapaz de llevar a cabo cierta acción aunque, en lo que atañe a estas consideraciones, esté en una

situación completamente adecuada para hacerlo. Asimismo, puede ser que la persona piense que debería llevar a cabo la acción y que tenga deseos de hacerlo. Si, no obstante, es incapaz de actuar porque encuentra que la acción es impensable, lo que le impide actuar es una cuestión volitiva. No puede llevar a cabo la acción porque *no puede tener la voluntad de llevarla a cabo.*

Alguien que ha decidido realizar cierta acción puede descubrir, a la hora de la verdad, que en realidad no puede hacerlo. En ese caso no está claro, estrictamente hablando, si en realidad tomó una decisión después de todo. Sin duda, se *decidió*, pero ello, obviamente, no logró configurar su *voluntad.* Desde luego, no obstante es posible insistir en que, en realidad, se tomó una decisión. Pero como de hecho no se logró ningún compromiso volitivo efectivo, insistir en que se tomó una decisión exigiría reconocer que la *decisión* no fue más que un acontecimiento meramente verbal o intelectual. Quizá sería mejor decir, en cuanto a situaciones problemáticas en este sentido, que no se tomó ninguna decisión *efectiva.*

Entonces, de una persona que no logra realizar una acción también se puede decir que no puede tomar una decisión efectiva de realizarla. La dificultad que esta persona tiene es que no puede organizarse de la manera necesaria desde el punto de vista volitivo. Al intentar hacerlo, tropieza con los límites de su voluntad. Esto lo demuestra el hecho de que es incapaz de realizar la acción incluso cuando se satisfacen todas las condiciones no volitivas para hacerlo (por ejemplo, la oportunidad y el poder). Quizá podría llevar a cabo la acción con facilidad si sólo tuviera la voluntad. Pero no puede desear realizar la acción. Para ella, tener la voluntad de realizarla es impensable.[5]

La categoría de lo impensable no es, en esencia, una categoría moral. Con seguridad, la incapacidad de actuar a veces puede derivar de consideraciones que son claramente morales. Se cuenta que algunos oficiales de las Fuerzas Armadas se negaron a llevar a

5 Existe una diferencia entre querer realizar una acción y tener la voluntad de hacerlo. Alguien que *tiene la voluntad* de realizar una acción al menos intentará realizarla. Pero tener un deseo o querer llevar a cabo cierta acción puede no conducir a la persona a ninguna acción cuyo motivo sea cumplir ese deseo. En ese caso, el deseo de la persona no es su voluntad.

cabo los procedimientos tendientes a lanzar armas nucleares cuando creyeron que las órdenes que habían recibido no eran parte de una prueba o un ejercicio y que la ejecución del procedimiento tendría como consecuencia un ataque nuclear real. Estos oficiales se habían ofrecido para la tarea; presumiblemente habían pensado que estaban dispuestos a cumplir órdenes como aquellas que desobedecieron. En el momento crítico, resultó que no pudieron actuar. Descubrieron que participar en el inicio de un ataque nuclear era, para ellos, impensable. Sin duda, ello se debió a inhibiciones morales. En otros casos, las consideraciones en virtud de las cuales algo es impensable pueden ser por completo dignas y no tener ningún significado moral. Es posible que alguien no logre realizar una acción porque se sentiría seriamente lastimado por ella, porque ésta dañaría su orgullo de modo intolerable o porque resulta demasiado desagradable.

El no lograr realizar una acción no es lo mismo que simplemente sentir una irresistible adversidad a realizarla. Por supuesto, la persona que no logra actuar tiene una aversión por la cual su conducta se ve restringida. Más aun, ella aprueba la aversión; y restringe su conducta de manera tan efectiva precisamente *debido a* ello.[6] El hecho de que la persona apruebe su aversión es lo que distingue las situaciones en las que alguien encuentra que una acción es impensable de aquellas en que su incapacidad de actuar se debe a una adicción o a algún otro tipo de impulso irresistible. Sin dudas es verdad que una persona que se ve limitada por un impulso irresistible también puede aprobar ese impulso. Sin embargo, en ese caso, la eficacia de la restricción que ese impulso le impone a su conducta no se debe a que la persona lo apruebe. Su aprobación del impulso no es lo que lo convierte en irresistible.

En el capítulo 59 de *The Eustace Diamonds*, Anthony Trollope describe a un hombre que no logra llevar adelante la conducta que, en forma libre y deliberada, había decidido seguir. Lord Fawn es un alto funcionario del gobierno y señor del reino. Invita a Andy Gowran, administrador de propiedades de humilde cuna y escasa instruc-

6 La naturaleza y la estructura de esta aprobación deben explicarse en detalle, pero no ahondaré aquí en el tema.

ción, para que le relate un incidente que él había presenciado: supuestamente, la prometida de Fawn (Lizzie Eustace) había abrazado a otro hombre (el primo de ella, Frank Greystock). Ahora Gowran, mientras expone su versión, le guiña un ojo a Fawn. A partir de ese momento, a Fawn le resulta imposible continuar con la entrevista:

> Fue terrible para Lord Fawn que el hombre le guiñara el ojo. No entendió exactamente lo último que había dicho Andy, pero sí entendió que este administrador escocés quería acusar a la mujer con la cual Fawn se había comprometido de cierta familiaridad indecente con el primo de ella. Cada sentimiento de su naturaleza se sublevó contra la tarea que tenía ante sí, y descubrió que, en la práctica, se le volvía absolutamente imposible. No podía investigar en detalle el coqueteo de la pobre Lizzie entre las rocas. Era débil y tonto, y en muchos aspectos, ignorante, pero era un caballero. A medida que se acercaba al tema que se suponía que había que tocar, más aborrecía a Andy Gowran y más se aborrecía a sí mismo por haber sucumbido a ese contacto. Hizo una pausa, y luego declaró que la conversación había llegado a su fin.

Para Lord Fawn es impensable compartir un asunto tan íntimo con una persona tan inferior a él en educación y clase. Había creído que sería una buena idea enterarse, a través de Gowran, de lo que Lizzie Eustace había estado haciendo exactamente con Frank Greystock, y se había decidido a hacerlo. Sin embargo, llegado el momento, "cada sentimiento de su naturaleza se sublevó". Se dio cuenta de que, simplemente, no podía hacerlo.

Los sentimientos de Lord Fawn no se sublevan contra su *voluntad*. Los sentimientos que para él tornan "absolutamente impracticable" continuar su entrevista con Gowran no se oponen a su voluntad. Se sublevan contra su intento de configurar su voluntad de cierto modo. Fawn piensa que lo mejor que puede hacer es hablar con Gowran acerca de Lizzie e intenta seguir ese curso de acción. Sin embargo, en el fondo, no está dispuesto a que su voluntad sea configurada de ese modo.

La voluntad de una persona puede ser dominada y violada por fuerzas, como las de la ansiedad o la adicción, que se generan den-

tro de ella, pero que, sin embargo, no son suyas en sentido pleno. Son fuerzas con las cuales no se identifica, cuya influencia se esfuerza por resistir. Cuando ellas dominan, la persona es dominada por ellas; no está en control de sí misma. Ahora bien, cuando alguien encuentra que una acción es impensable para él, lo que descubre no es que sea incapaz de mantener el control de sí mismo. Ello supondría que el control ha pasado a fuerzas que no son verdaderamente las suyas. No obstante, en un caso como el de Fawn, las fuerzas que se sublevan y toman el control –es decir, "todo sentimiento de su naturaleza"– no se oponen al agente. Son, en el sentido más auténtico, sus propias fuerzas, que no sólo se encuentran en su interior simple y literalmente, sino que también están integradas en su naturaleza. Aunque le impiden llevar a cabo una acción que había creído que quería realizar, lo hacen sólo debido al hecho de que él, en realidad, no quiere llevarla a cabo. Su incapacidad para cumplir con la acción revela que él no tiene la voluntad de desearla.

Una persona que afirma que considera impensable una acción quiere decir que no existen circunstancias en las cuales ella estaría dispuesta a llevarla a cabo. Esto puede reflejar cierta escasez de imaginación de su parte. Quizá realmente haya circunstancias posibles que no ha considerado o evaluado, en las que realizaría la acción con muy buena voluntad. Sea como fuere, su afirmación de que la acción le resulta impensable no le exige sostener que no podría haber circunstancias en las que ella sería su mejor alternativa. Tampoco le exige negar que podría haber circunstancias en las que sería razonable que él llevara a cabo la acción. Después de todo, es muy posible que alguien anticipe en forma realista que sería incapaz de seguir cierto curso de acción incluso si reconociera que es el mejor. Para él, puede ser evidente que hay cuestiones respecto de las cuales es incapaz de actuar de modo racional.

IV

Existe un conocido modo de interpretar la noción de *racionalidad* que, a pesar de su importancia, los filósofos a menudo descuidan.

Entre aquellos que lo hacen, se destaca Hume. Él insiste en que un deseo o una preferencia son contrarios a la razón sólo si se basan en un juicio falso ya sea acerca de lo que existe, ya acerca de alguna relación causal. Este modo restringido de comprender la racionalidad conduce a Hume a hacer algunas asombrosas afirmaciones:

> Cuando la pasión no se funda en falsos supuestos ni escoge medios insuficientes para su fin, el entendimiento no puede justificarla ni condenarla. No es contrario a la razón preferir la destrucción del mundo entero a un arañazo en mi dedo. No es contrario a la razón que yo prefiera arruinarme totalmente con el fin de evitarle la menor preocupación a un indio o a una persona totalmente desconocida. Es tan poco contrario a la razón preferir incluso un bien que reconozco como menor al bien mayor y experimentar una afección más ardiente por el primero que por el último. [...] En resumen, una pasión debe ir acompañada por algún juicio falso para ser irracional, y aun así, lo que es irracional, hablando con propiedad, no es la pasión, sino el juicio.[7]

Las preferencias como las que aquí describe Hume pueden ser independientes de cualquier error fáctico. Así, la facultad de la razón o del entendimiento no pueden ni justificarlas ni condenarlas si, como él sostiene, a esa facultad sólo le atañe distinguir entre juicios fácticos verdaderos y falsos. En este sentido, las preferencias, por más exorbitantes o extrañas que puedan ser, no son irrazonables.

No obstante, en un sentido bien establecido y valioso, las pasiones y las preferencias a las que se refiere Hume son, categóricamente, no razonables. Consideremos qué deberíamos decir acerca de alguien a quien en realidad le importara más la herida en su dedo que la destrucción del mundo entero. Diríamos que *debe de estar loco*. Dicho de otra manera, le atribuiríamos un defecto de la razón. Más aun, esta atribución de irracionalidad estaría justificada. Diga lo que dijere Hume, considerar que la destrucción del

7 *A treatise of human nature*, ed. por L. A. Selby-Bigge, Oxford, Oxford University Press, 1888, Libro II, Parte III, Sección III, p. 416 [trad. esp.: *Tratado de la naturaleza humana*, Madrid, Tecnos, 1988].

mundo entero es menos importante que un dedo rasguñado no es una opción racional. Es *lunática.*

Por consiguiente, debe de haber otra manera, distinta de la de Hume, de comprender qué significa ser contrario a la razón. De no ser así, no podríamos emplear con precisión términos como *lunático, loco* o *irracional* para calificar las preferencias en cuestión, a menos que usáramos esos términos de manera muy laxa como vehículos de un abuso meramente genérico. A mi criterio, cuando decimos que alguien con una preferencia de este tipo debe de estar loco, no pretendemos sólo expresar una desaprobación general y vaga. Nuestra intención específica es acusar a la persona de un tipo de irracionalidad. Desde luego, esto no significa que no la acusaríamos también de ser malvada. No obstante, el problema que ella tiene no puede expresarse con exactitud en términos que simplemente atañen a su carácter moral. Sería sumamente insuficiente decir que su problema es que su carácter moral es pobre, y esto sería insuficiente incluso si aumentáramos la intensidad de la afirmación diciendo que es *extraordinariamente* malvada o que su carácter moral es *extremadamente* pobre. No resulta satisfactorio interpretar que su maldad es una falla exclusivamente moral.

La inmoralidad no es, como tal, anormal. En muchos casos, es fácil reconocer que deriva de la naturaleza humana de un modo bastante corriente. No obstante, la inmoralidad del ejemplo en cuestión es de un orden especial. No entra dentro del rango de lo que entendemos como normal; en cambio, nos resulta antinatural o, incluso, monstruosa. Por ello, las nociones de irracionalidad y de demencia parecen guardar relación. La racionalidad pertenece de manera distintiva a la naturaleza esencial de un ser humano. Si consideramos que un juicio o una elección se oponen a la naturaleza humana –es decir, si nos resultan antinaturales o inhumanos–, nos sentimos inclinados a pensar que, por tanto, conllevan un defecto de la razón.[8]

No obstante, ¿qué significan en realidad todos estos epítetos? ¿Cuál es el carácter específico de la irracionalidad al cual aluden?

8 Los términos *demente, antinatural* e *irracional* son convergentes. Los dementes son irracionales, y la irracionalidad es antinatural en una criatura para cuya naturaleza la razón es esencial.

Hume atribuye *irracionalidad* sólo cuando una creencia no es coherente consigo misma o con los hechos. Alguien que prefiere su dedo al mundo perfectamente podría no ser irracional en este sentido. Su defecto es volitivo, y no epistémico o cognoscitivo. La persona es anormal no sólo por lo que cree, sino por lo que es capaz de hacer. La razón por la que la consideramos irracional es que no le parece que cierta conducta sea impensable.

Desde luego, lo que justifica la atribución de irracionalidad no puede ser meramente que la persona sea capaz de una conducta que *nosotros* consideramos impensable. La mayoría de nosotros, simplemente por el hecho de que seríamos incapaces de llevar a cabo cierta acción, suponemos que esta incapacidad que tenemos es una marca o un requisito general y confiable de la cordura. No hace falta decir que alguien que es capaz de hacer lo que nosotros consideramos impensable puede carecer de defectos. Bien puede ser, después de todo, que nuestra propia incapacidad sea anormal. O puede ser que no haya ninguna anormalidad ni en el hecho de poder ni en el de no poder llevar a cabo la acción en cuestión. Quizá tanto la capacidad como la incapacidad dependan de creencias, compromisos o idiosincrasias personales –relacionadas con cuestiones de gusto, política, religión, amor u otras– respecto de los cuales resulta natural y de ninguna manera patológico que las personas tengan opiniones diversas.

Con seguridad, a veces consideramos que lo que creemos impensable define un criterio de normalidad. La cualidad de impensable no nos parece ser meramente personal, sino que, por alguna razón, tiene una trascendencia más general. Sin embargo, incluso en aquellos casos es probable que reconozcamos que el alcance del criterio es limitado. Sabemos que las preferencias o los tipos de conducta que son irracionales en un ámbito cultural a menudo pueden ser por completo racionales en otro. Consideremos, por ejemplo, este relato:

> Uno de los últimos integrantes de la familia de payasos Fratellini, un hombre mayor, hizo hace algunos años unas declaraciones para la televisión de París, en las que [explicó] por qué escaseaban los buenos payasos de circo jóvenes. "Cuando era niño,

> mi padre, bendito sea, me rompió las piernas para que caminara de un modo cómico, tal como debería hacerlo un payaso", dijo el anciano. [...] "Ahora hay personas que no verían con buenos ojos ese tipo de cosas".[9]

Lo que para algunas personas es impensable, para otras puede ser no sólo perfectamente razonable, sino también exquisitamente correcto.

Más aun, lo que es impensable para una persona puede variar de una época a otra. Las necesidades de la voluntad no son –necesariamente ni siempre– permanentes. Están sujetas a cambios, según los cambios en las circunstancias contingentes de las que derivan. Es incluso posible que una persona desee que los límites de su voluntad sean diferentes de lo que son y que desee alterarlos de manera deliberada. Supongamos que una acción (por ejemplo, comer carne humana) le repugna tanto a la persona que no puede hacerla. Pero ahora supongamos que se encuentra en circunstancias que la llevan a reconocer que esa incapacidad es desfavorable para sus intereses (quizá no tenga nada para comer, salvo personas) y que no tiene una razón preponderante para desear conservarla (nunca se trató de un principio meditado, sino sólo de una reacción visceral). Entonces, podría tomar medidas para alterar su capacidad volitiva a fin de ser capaz de hacer lo que ahora considera impensable.

No hace falta decir que para la persona sería imposible cambiar de ese modo por un mero acto de la voluntad. Ella no podría alterar su voluntad sólo decidiendo que lo que le resultaba impensable ya no lo es más. Ello podría ser posible si la tarea consistiera sólo en superar una inhibición poderosa; a veces una inhibición puede superarse mediante un poderoso esfuerzo de la voluntad solamente. Pero una voluntad limitada por la necesidad genuina difícilmente podría ser alterada mediante ese tipo de esfuerzo. Sin embargo, alguien a quien la acción le resulta impensable tiene, seguramente, la posibilidad de intentar –por otros medios menos directos que el solo ejercicio de la fuerza de la voluntad– alterar su

9 A. J. Liebling, *Between meals*, Nueva York, Simon and Schuster, 1962, p. 149.

propia voluntad de manera tal que la acción se torne pensable para él. El hecho de que una persona no logre realizar una acción no implica que no pueda actuar con la intención de cambiar ese hecho. Ni tampoco implica que no pueda tener éxito en la concreción de su intención.

No obstante, en ciertos casos, actuar con la intención de tornar pensable lo impensable es, en sí mismo, algo que la persona no logra hacer. No sólo puede resultarle impensable llevar a cabo cierta acción, también puede resultarle impensable formar una intención efectiva para sentirse dispuesta a realizarla. No logra aprobar la idea misma de cambiar de ese modo. Así, el carácter impensable de la acción es tan decisivo que para la persona constituye un límite no sólo para lo que puede hacer, sino también para lo que puede ser. Es una necesidad genuina de su naturaleza volitiva. Desde luego, ésta puede cambiar, ya que es una cuestión lógicamente contingente y, por tanto, susceptible de ser afectada por fuerzas causales. Sin embargo, la persona misma no puede cambiarla deliberadamente. De este modo, al menos en parte, la necesidad define lo que debe ser la persona. En este sentido, es un elemento constitutivo de su naturaleza o de su esencia como persona.

Los triángulos tienen alternativas a la de ser isósceles. El hecho de ser isósceles es, por tanto, una característica accesoria o accidental de un triángulo. En cambio, el hecho de que los ángulos interiores de cualquier triángulo sumen 180° es una característica esencial. Un triángulo no tiene más alternativa que tener esta característica; en ninguna circunstancia un triángulo tendrá ángulos interiores que sumen más o menos de 180°. De manera similar, en ninguna circunstancia una persona querrá hacer lo que para ella es impensable. Así como el conjunto de las características esenciales de un triángulo especifica los límites de lo que éste puede ser, el conjunto de las acciones que a una persona le resultan impensables especifica los límites de lo que ésta puede tener la voluntad de hacer. Define su esencia como criatura volitiva. Hay individuos que están dispuestos a hacer *lo que sea* si las consecuencias son lo suficientemente deseables, es decir, si el precio es el correcto. Lo que tienen la voluntad de hacer, por tanto, nunca está determinado exclusivamente por su propia naturaleza, sino que siempre es una

función de sus circunstancias. Esto significa que en lo que concierne a sus voluntades, sólo tienen características accidentales. Y en la medida en que una persona es una entidad volitiva, el individuo es una persona sin ninguna naturaleza esencial.

V

Por lo general, se presume que una persona actúa bajo la guía de la razón y que tiene el control de sí misma sólo cuando lo que hace concuerda con lo que juzga que debe hacer. Si sus sentimientos vencen o reemplazan su juicio, se supone que ha perdido su autocontrol racional. Según esta manera de mirar las cosas, Lord Fawn pierde su autocontrol y deja de ser guiado por la razón cuando descubre que no logra continuar su conversación con Andy Gowran. Supongamos que decimos que ser guiados por la razón es actuar en conformidad con el propio juicio. Entonces, Fawn no está siendo guiado por la razón cuando la rebelión de sus sentimientos arrasa con el juicio que ha hecho respecto del mejor curso de acción. Más aun, en la medida en que somos proclives a identificar a una persona con lo que piensa, no consideraremos que Fawn tiene el dominio de sí mismo cuando, en cambio, es dominado por lo que siente.

No obstante, me parece que esta manera de mirar las cosas es errónea. Es un error fundamental considerar que cualquier oleada de emoción contra el juicio es un levantamiento de lo irracional. Con seguridad, hay un sentido algo trivial en que los sentimientos son intrínsecamente no racionales: no se relacionan con la facultad de la razón, porque no son en esencia discursivos. En un sentido más sustancial, los sentimientos pueden concordar mejor con la razón que el juicio. El juicio de una persona puede ser en sí mismo radicalmente contrario a la razón. Por tanto, el hecho de que su juicio guíe su conducta no significa, en sí mismo, que esté actuando racionalmente. En realidad, bien podría ser que el desajuste entre su voluntad y su juicio sea precisamente lo que la salva de la irracionalidad. Existe –no sólo entre los filósofos, sino

también en el ámbito del derecho e, incluso, en el sentido común– una desafortunada tendencia a suponer que cuando la voluntad de una persona está tan poderosamente restringida por sus emociones que no puede evitar actuar de cierta manera, su situación en este aspecto es patológica. No obstante, de hecho puede ocurrir justamente lo contrario.

Supongamos que un individuo sólo tiene dos opciones: puede sufrir una herida menor en el dedo o puede causar la destrucción del mundo. ¿Es una condición necesaria de la cordura que, en lo que se refiere a la influencia de sus emociones, sea tan capaz de elegir la última alternativa como de elegir la primera? ¿Acaso el equilibrio emocional que supone la salud mental requiere que la persona posea la misma capacidad para elegir uno u otro curso de acción? La conducta de una persona psicológicamente sana ¿debe depender sólo de su juicio respecto de cuál de las dos alternativas es preferible? Con seguridad, no es así. Alguien que decidió sacrificar el mundo con el objetivo de salvar su dedo daría, de esa manera, un indicio convincente de un serio trastorno mental. No obstante, el indicio sería considerablemente más grave si no sólo hiciera este juicio, sino que también demostrara ser capaz de concretarlo efectivamente. En ese caso, sin dudas estaría completamente loco.

Por otra parte, supongamos que alguien decidió permitir que se destruyera el mundo por el bien de su dedo, pero no pudo seguir adelante con su intención porque una irresistible ola de emoción le impidió, a su pesar, que lo hiciera. En otras palabras, supongamos que sus sentimientos se sublevaron contra su juicio y que le impidieron hacer lo que había decidido hacer deliberadamente. La emergencia de su incapacidad de realizar la acción en cuestión *reivindicaría* su cordura sustancialmente, o, al menos, en cierta medida. Demostraría que la irracionalidad manifiesta en su horrendo juicio no era, después de todo, muy profunda.

La cualidad de impensable es un modo de necesidad al que, a veces, la voluntad se une y que limita las elecciones. Esta limitación puede ser una afirmación y una revelación de una cordura esencial. Hay ciertas cosas que ningún individuo completamente racional consideraría hacer nunca. Pero si alguien de alguna manera *consi-*

derara hacerlas y llegara hasta el punto de *decidirse* a hacerlas, una persona básicamente cuerda no podría lograr *hacerlas* realmente. La cordura consiste en parte en ser susceptible, justamente, a esas incapacidades.

Con frecuencia se presume erróneamente que las inclinaciones irresistibles y las inhibiciones insuperables son, invariablemente, síntomas o constituyentes de un trastorno mental o emocional. A veces son, en realidad, manifestaciones de la racionalidad esencial de una persona, una racionalidad tan profundamente arraigada que no puede volverse impotente por una mera convicción o una intención fugaz. Esta racionalidad esencial puede, a veces, proteger a la persona, movilizando sus emociones, de sucumbir a la influencia de alteraciones radicales de su juicio.

Por tanto, la voluntad de un agente racional no tiene por qué estar desprovista de carácter sustancial. No necesariamente es por completo formal y carente de satisfacciones, sin predisposiciones inherentes propias. Si la voluntad de una persona fuera un instrumento por completo carente de características, sin otra capacidad que la de transmutar su juicio sobre qué hacer en una expresión efectiva de sus poderes activos, se asemejaría mucho a la *persona desnuda* a la que, según Rawls, el utilitarismo reduce al agente de la elección racional. De hecho, sin embargo, precisamente en el contenido particular o en el carácter específico de su voluntad –que saludablemente puede conducirlo a actuar en contra de su juicio– puede residir, en parte, la racionalidad de una persona.

Existe un modo de racionalidad que concierne a la voluntad misma. Al igual que el modo de racionalidad que se expresa en las verdades necesarias de la lógica, tiene que ver con la inviolabilidad de ciertos límites. Las necesidades lógicas definen lo que para nosotros es imposible concebir. Las necesidades de la voluntad atañen a lo que somos incapaces de hacer.

Fuentes

El último ensayo de esta recopilación se publica por primera vez. Los demás, a los que se les han introducido cambios estilísticos menores, ya habían aparecido en otras publicaciones, en el orden en que se presentan en los capítulos.

1. *Journal of Philosophy*, LXVI, N° 23, 4 de diciembre de 1969.
2. *Journal of Philosophy*, LXVIII, N° 1, 14 de enero de 1971.
3. *Essays on freedom of action*, ed. por Ted Honderich, Londres, Routledge & Kegan Paul, 1973.
4. *Proceedings of the aristotelian society*, volumen complementario, 1975.
5. "Identification and externality" se publicó originalmente en *The identities of persons*, editado por Amelie Rorty, University of California Press, 1977. Copyright: © 1976, The Regents of the University of California.
6. *American Philosophical Quarterly*, 15, 1978.
7. *Synthese*, vol. 53, N° 2, 1982, pp. 257-272. Copyright: © 1982, D. Reidel Publishing Company.
8. *How many questions? Essays in honor of Sidney Morgenbesser*, ed. por L. S. Cauman, Isaac Levi, Charles D. Parsons, Robert Schwartz, Indianápolis, Hackett Publishing Co., 1983.
9. *Philosophy and phenomenological research*, XLV, N° 1, 1984.
10. *Raritan*, VI, N° 2, 1986.
11. *Ethics*, vol. 98, N° 1, octubre de 1987.
12. *Responsibility, character, and the emotions: New essays in moral psychology*, ed. por Ferdinand David Schoeman, Nueva York, Cambridge University Press, 1987. Copyright: © 1987, Cambridge University Press.

Este libro se terminó de imprimir
en diciembre de 2005 en Latingráfica S.R.L.
(www.latingrafica.com.ar), Rocamora 4161
CP C1184 ABC, Buenos Aires.

Los contenidos de este libro pueden ser reproducidos, en todo o en parte, siempre y cuando se cite la fuente y se haga con fines académicos, y no comerciales

Otros títulos

cn 3006 *Cass R. Sunstein*
Riesgo y razón
Seguridad, ley y medioambiente

cn 3008 *Martha C. Nussbaum*
El ocultamiento de lo humano
Repugnancia, vergüenza y ley

En preparación

cn 3014 *Harry G. Frankfurt*
Necesidad, voluntad y amor

df 1006 *Odo Marquard*
Felicidad en la infelicidad

cn 3015 *Heinrich Wiegand Petzet*
Encuentros y diálogos con Martin Heidegger, 1929-1976

Este libro reúne trece ensayos seminales sobre ética y filosofía de la mente escritos por uno de los más importantes filósofos norteamericanos contemporáneos. Los trabajos, que se ocupan de problemas fundamentales como la libertad de la voluntad, la naturaleza de la acción, la constitución del yo y la teoría de los ideales de las personas, están sin embargo atravesados por algunos temas recurrentes: el papel esencial de la necesidad para el logro de una autonomía genuina, la naturaleza y las condiciones de la responsabilidad moral, los modos en que nos conceptualizamos a nosotros mismos como personas. El libro comienza con el ensayo titulado "Posibilidades alternativas y responsabilidad moral", sin dudas el trabajo más influyente de la segunda mitad del siglo XX referido a la responsabilidad moral. Los ensayos siguientes inician una serie de debates acerca de la autonomía humana y la estructura del yo, cuestiones que continúan dominando la literatura sobre la materia. El volumen incluye la versión original del famoso ensayo "On bullshit", que, bajo la forma de libro independiente, encabezó, durante varios meses del año 2005, la lista de libros más vendidos de no ficción en el *New York Times*. En nuestro sello editorial aparecerá, próximamente, otra obra fundamental de Harry Frankfurt: *Necesidad, voluntad y amor*.

"En el mundo académico, los textos escritos con claridad son antes la excepción que la regla. Por ello es tan refrescante leer el trabajo de Harry Frankfurt. Sus análisis acerca de las motivaciones de las personas son extraordinariamente claros y directos." **Bart Pattyn**, "Ethical perspectives", en *Journal of the European Ethics Network* 5, 1

katz www.katzeditores.com

ACCESO ABIERTO

isbn 84-609-8358-7

9 788460 983583